DOCUMENT RESUME

ED 061 682 EC 041 770

AUTHOR	Foulke, Emerson, Ed.
TITLE	Proceedings of the Second Louisville Conference on Rate and/or Frequency-Controlled Speech (Louisville, Kentucky, October 22-24, 1969).
INSTITUTION	Louisville Univ., Ky. Perceptual Alternatives Laboratory.
PUB DATE	Feb 71
NOTE	414p.; Prepared at the Center for Rate-Controlled Recordings, Perceptual Alternatives Laboratory, University of Louisville
EDRS PRICE	MF-$0.65 HC-$16.45
DESCRIPTORS	Conference Reports; Handicapped Children; *Listening Comprehension; Research Projects; *Speech Compression; *Time Factors (Learning)

ABSTRACT

Proceedings of the conference on rate and/or frequency-controlled speech contain 33 conference papers. Reports fall into three basic categories: basic research concerning the perception and listening comprehension of time and/or frequency-controlled speech; technical reports concerning the production of time and/or frequency-controlled speech; and reports of practical applications of such speech in educational, industrial, and other settings. Applications reported in educational settings involve both normal and handicapped (blind, speech handicapped, hard of hearing) subjects. (KW)

ED 061682

PROCEEDINGS OF THE
SECOND LOUISVILLE CONFERENCE ON
RATE AND/OR FREQUENCY-CONTROLLED SPEECH

October 22-24, 1969

Dr. Emerson Foulke, Editor

February 1971

CENTER FOR RATE-CONTROLLED RECORDINGS
A Unit of the Perceptual Alternatives Laboratory

UNIVERSITY OF LOUISVILLE

LOUISVILLE, KENTUCKY 40208

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION
Emerson Foulke 1

II. AN INTRODUCTION TO SPEECH TIME COMPRESSION TECHNIQUES: THE EARLY DEVELOPMENT OF SPEECH TIME COMPRESSION CONCEPT AND TECHNOLOGY
H. Leslie Cramer 6

III. EFFECT OF RATE OF COMPRESSION AND MODE OF PRESENTATION ON THE COMPREHENSION OF A RECORDED COMMUNICATION TO JUNIOR COLLEGE STUDENTS OF VARYING APTITUDES
Clement Cordell Parker 21

IV. PERTURBATIONS OF SEX JUDGMENTS WITH TIME-COMPRESSED AND FREQUENCY-ALTERED SPEECH
Daniel S. Beasley and Willard R. Zemlin 29

V. DICHOTIC SPEECH-TIME COMPRESSION
Sanford E. Gerber and Robert J. Scott 47

VI. A COMPARISON OF "DICHOTIC" SPEECH AND SPEECH COMPRESSED BY THE ELECTRO-MECHANICAL SAMPLING METHOD
Emerson Foulke and E. McLean Wirth 60

VII. THE INTELLIGIBILITY OF COMPRESSED WORDS
Robert Heise 70

VII. STUDIES ON THE EFFICIENCY OF LEARNING BY LISTENING TO TIME-COMPRESSED SPEECH
Thomas G. Sticht 80

IX. THE APPLICATION OF RATE-CONTROLLED RECORDINGS IN THE CLASSROOM
Richard W. Woodcock 90

X. PROCESSING TIME AS A VARIABLE IN THE COMPREHENSION OF TIME-COMPRESSED SPEECH
Ruth Ann Overmann 103

XI. SPEECH PAUSE DURATION AS A FUNCTION OF SYNTACTIC JUNCTURES
Kenneth F. Ruder and Paul J. Jensen 119

XII. THE SIGNIFICANCE OF INTRA- AND INTER-SENTENCE PAUSE TIMES IN PERCEPTUAL JUDGMENTS OF ORAL READING RATE
Norman J. Lass 134

XIII. EFFECTS OF TEMPORAL SPACING ON LISTENING COMPREHENSION: A SOURCE OF INDIVIDUAL DIFFERENCES
Raymond L. Johnson and Herbert L. Friedman 143

XIV. THE COMPREHENSION OF RATE INCREMENTED AURAL CODING
Murray S. Miron and Eric Brown 153

XV. THE EFFECTS OF SOURCE LANGUAGE RATE ON THE PERFORMANCE OF SIMULTANEOUS CONFERENCE INTERPRETERS
David Gerver 162

XVI. THE A. F. B. HARMONIC COMPRESSOR
John W. Breuel and Leo M. Levens 185

XVII. THE GRAHAM COMPRESSOR, A TECHNICAL DEVELOPMENT OF THE FAIRBANKS METHOD
Wayne Graham 193

XVIII. TIME COMPRESSION OF SPEECH ON A SMALL COMPUTER
S. U. H. Qureshi and Y. J. Kingma 196

XIX. THE BRAIDED-SPEECH METHOD OF TIME COMPRESSING SPEECH
H. Leslie Cramer and Robert P. Talambiras 206

XX. THE MEASUREMENT OF LISTENING COMPREHENSION
David B. Orr 219

XXI. EFFECTS OF MOTIVATION AND WORD RATE ON AURAL COMPREHENSION
Carson Y. Nolan and June E. Morris 224

XXII. COMPREHENSION OF NARRATIVE PASSAGES BY FOURTH-GRADE CHILDREN AS A FUNCTION OF LISTENING RATE AND ELEVEN PREDICTOR VARIABLES
Robert L. Gropper 231

XXIII. A COMPARISON OF TWO TECHNIQUES FOR INCREASING THE READING RATE OF SIXTH-GRADE GIFTED PUPILS: THE COMPRESSED SPEECH MACHINE AND THE SELF-IMPROVEMENT METHODS
Charles R. Walker 253

XXIV. USING COMPRESSED SPEECH TO TEACH INSTRUCTIONAL TECHNIQUES TO AIR FORCE OFFICERS
Meredith W. Watts, Jr. 260

XXV. ECONOMIC ANALYSIS OF TIME-COMPRESSED SPEECH FOR INSTRUCTIONAL BROADCAST SATELLITES: A PROPOSAL FOR BRAZIL
D. T. Jamison 271

XXVI. AN INVESTIGATION INTO EXTENDED USE OF TIME-COMPRESSED SPEECH WITH INTERMEDIATE-GRADE SUBJECTS
Grace D. Napier 282

XXVII. PROGRAM OF TRAINING FOR LISTENING-READING FOR VISUALLY IMPAIRED STUDENTS USING COMPRESSED SPEECH RECORDINGS
Rachel Rawls 289

XXVIII. COMPREHENSION FOR IMMEDIATE RECALL OF TIME-COMPRESSED SPEECH AS A FUNCTION OF SEX AND LEVEL OF ACTIVATION
Sally McCracken 313

XXIX. THE COMPREHENSION OF RATE-CONTROLLED SPEECH BY SECOND-GRADE CHILDREN WITH FUNCTIONAL MISARTICULATIONS
R. Vernon Stroud 320

XXX. RATE-CONTROLLED SPEECH AND SECOND LANGUAGE LEARNING
Herbert L. Friedman and Raymond L. Johnson 323

XXXI. COMPRESSED SPEECH IN MEDICAL EDUCATION
Gloria J. Boyle 328

XXXII. THE RELATIONSHIP OF LISTENING SKILLS TO THE UTILIZATION OF COMPRESSED SPEECH
Rolland Callaway, Gerald Gleason, and Barbara Klaeser 332

XXXIII. DEAF CHILDREN'S AUDITION OF DISTINCTIVE FEATURES WITHIN FREQUENCY SAMPLED SPEECH
Daniel Ling 338

XXXIV. THE EFFECTS OF TRAINING ON THE INTELLIGIBILITY AND COMPREHENSION OF FREQUENCY-SHIFTED TIME-COMPRESSED SPEECH BY THE BLIND
Paul E. Resta 347

LIST OF REFERENCES 363

CHAPTER I

INTRODUCTION

Emerson Foulke*

Of all of the forms of communication in which humans engage, perhaps the most important is that which depends upon the interaction of one or more speakers and one or more listeners. However, we have tended to take it for granted, and we have not subjected it to the intensive scrutiny given to other forms of communication, such as written communication.

Throughout most of man's history, proximity has been a necessary condition for communication between speakers and listeners. However, because of the radio and the telephone, both of which have been developed largely within this century, spatial proximity is no longer a necessary condition and, because human speech can now be recorded for subsequent reproduction, temporal proximity is no longer a necessary condition either.

Until recently, there has been no way to gain significant control over the rate of communication between speakers and listeners. This rate has been determined primarily by the cognitive and articulatory limitations of speakers, and has not been amenable to the preferences or capabilities of listeners. However, with the advent of methods such as the one described by Grant Fairbanks and his co-workers at the University of Illinois (Fairbanks, Everitt, & Jaeger, 1954), it has become possible to vary the rate of recorded speech without materially affecting its other parameters.

In the past few decades, there has been a growing awareness of the educational importance of the communication that takes place between speakers and listeners, and of the possibilities afforded by modern communication technology for increasing its flexibility and efficiency as an educational tool. One manifestation of this new awareness is the growing number of courses offered in educational and industrial settings for the purpose of improving listening skills. A special interest has been expressed by those concerned with the education of children who, for whatever reason, must place extraordinary reliance on speaking and listening in order to communicate. Blind school children, for instance, depend heavily upon reading by listening because they cannot read print and because they read braille so slowly. There is an increasing awareness on the part of educators of the large number of children without visual impairment who have serious reading problems that do not yield to remedial efforts, and the advantage they might gain from reading by listening is beginning to receive attention. Much of the instruction provided by college and industry depends upon aural communication, and the feasibility of increasing the rate of recorded speech suggests intriguing possibilities for more efficient utilization of the limited time available for instruction.

---

*Dr. Emerson Foulke, Director of the Perceptual Alternatives Laboratory, University of Louisville, is Editor of the Proceedings of the Second Louisville Conference on Rate and/or Frequency-Controlled Speech.

The ability to reduce the rate of recorded speech may also be valuable. Word rates that are slower than normal may, in some cases, be more compatible with the cognitive abilities of mentally retarded children. Students of foreign language and individuals with problems of articulation may profit by the opportunity to hear the phonetic components of spoken words at a slower rate. Recorded speech presented at a rate that is slower than normal may afford a technique for pacing slow readers or students of typing. Secretaries may be able to transcribe recorded dictation more efficiently when they listen to speech, the word rate of which has been reduced.

The first method for altering the word rate of recorded speech to receive the attention of investigators (Fletcher, 1929; Klumpp & Webster, 1961) was the reproduction of a tape or record at a different speed than the one used during recording. This method achieves the desired effect as far as word rate is concerned. Reproduction at a faster speed increases word rate, while reproduction at a slower speed reduces word rate. However, in either case, serious distortion is introduced that soon renders words unintelligible. Fortunately, the method developed by Fairbanks et al. is now available. This is a sampling method in which periodic samples of a recorded signal are reproduced in order and with temporal contiguity. If the duration of the samples discarded by this procedure is brief enough, the ear will not be able to detect their absence, and if the time required for the reproduction of each critical feature of a speech signal is greater than the time represented by each discarded sample, it will be impossible for the critical feature to fall entirely within a discarded sample. These conditions are satisfactorily met when discarded samples are 30 milliseconds (msec.) in duration or less, and the result is speech, the word rate of which has been increased without distortion in pitch or voice quality. A recorded signal may be expanded in time by reproducing overlapping samples of that signal and the result is a reduction in word rate without pitch distortion.

The control of speech rate that was made possible by the commercial availability of equipment for sampling speech in the manner just described has stimulated a great deal of research concerning the effect of speech rate on the intelligibility of words and phrases, and the comprehensibility of fluent speech (Fairbanks, Guttman, & Miron, 1957c; Fairbanks & Kodman, 1957; Foulke, Amster, Nolan, & Bixler, 1962; Foulke & Sticht, 1967a; Friedman, Orr, Freedle, & Norris, 1966; Garvey, 1953b; Orr & Friedman, 1964). Experimental attention has been given to a variety of questions in which word rate figures as a factor. There has been an accumulation of experimental results which support the general conclusion that speech may be presented at a rate in the neighborhood of 275 words per minute (wpm) with the expectation of satisfactory comprehension, and if an appropriate training experience can be devised, comprehension of speech at much higher word rates may be possible as well. Because of these findings, many people have begun to give serious consideration to the benefits that might be realized by the use of rate-controlled recorded speech and there has been a steadily increasing interaction between researchers and educators in developing its practical applications. In addition, those interested in basic research on the perception of speech have taken advantage of the opportunity to control speech rate while holding other parameters constant (Foulke & Sticht, 1967a; Friedman & Johnson, 1968; Miron & Brown, 1968; Overmann, 1969; Wilson, 1969).

The first Louisville Conference on Time-Compressed Speech was convened at the University of Louisville on October 19, 20, and 21, 1966. The Conference was presented under the joint sponsorship of the Library of Congress and the University of Louisville, with additional financial support from the Office of Education. A volume containing the proceedings of the Conference, and an extensive list of references to the research literature on rate-controlled recorded speech, was prepared and distributed. This volume has proved to be a valuable source of information for those interested in rate-controlled recorded speech.

Another outcome of the Conference was the appointment of an implementation committee, charged with the responsibility of promoting action on recommendations developed during the Conference. One of the most urgent recommendations of the Conference was the establishment of a national center from which rate-controlled speech could be obtained. In response to this recommendation, the Center for Rate-Controlled Recordings was established at the University of Louisville, under the direction of Dr. Emerson Foulke, with the implementation committee serving as its Advisory Board. Since that time, the Board has met two or three times each year to discuss the development of rate-controlled recorded speech as a tool for research and education, to review the activities of the Center, and to participate in the formulation of new Center projects.

Another urgent recommendation of the first Louisville conference was for the development of a mechanism for disseminating information about rate-controlled recorded speech to those interested in its applications. In response to this suggestion, the Center undertook the publication of a monthly newsletter which reports research plans and findings, new applications, equipment development, and other information of interest to workers in the field. In addition, the Center fills requests for research reports and demonstration tapes containing samples of recorded speech, compressed or expanded in time by the several known methods.

Since the first Louisville conference, there has been a rapid growth in the level of interest and activity concerning rate-controlled recorded speech. Accordingly, the Center's Board decided to convene a second Louisville conference to serve this interest and the related interest of frequency-controlled speech. The Second Louisville Conference on Rate and/or Frequency-Controlled Speech was held at the University of Louisville on October 22, 23, and 24, 1969, under the sponsorship of the University of Louisville, with financial support from the American Foundation for the Blind, the Library of Congress, and the Office of Education. This Conference was attended by approximately 125 people, representing such fields as psychology, linguistics, education, educational administration, library science, engineering, and industry. The Conference program consisted of reports in three categories: basic research concerning the perception of time and/or frequency-controlled speech; technical reports concerning the production of time and/or frequency-controlled speech; reports of practical applications of such speech in educational, industrial, and other settings. A preconference workshop was held for the purpose of providing some exposure to relevant terms and concepts for those unfamiliar with the area. The first conference day included a luncheon meeting with Dr. A. Hood Roberts as the guest speaker.*

This volume contains the 33 conference reports. Since there was considerable overlap in the references cited by authors, it was decided not to include a list of references at the end of each report. Instead, the references cited by authors have been combined into a single list. This list has been augmented by entries from the reference file maintained by the Center for Rate-Controlled Recordings, and from a list of references prepared by Dr. Daniel S. Beasley, Department of Audiology and Speech Sciences, Michigan State University, and Dr. Willard R. Zemlin, Voice and Hearing Sciences Research Laboratory, University of Illinois. This list of references, though possibly not bibliographic in scope, is extensive, and it is hoped that it will serve as a valuable resource to those wishing to read in the area.

---

*Dr. A. Hood Roberts is affiliated with the Center for Applied Linguistics, Washington, D.C. The title of his luncheon address was "Automation and Speech."

In some cases, Conference reports were written by more than one author. Unless otherwise indicated, these reports were presented by the senior authors. Dr. Daniel Ling and Dr. Paul Resta were scheduled to make reports to the Conference. Due to circumstances beyond their control, they were unable to attend the Conference. Nevertheless, their reports have been included in this volume.

Mr. Stephen F. Temmer, President of Infotronic Systems, Inc., reported on the Information Rate Changer, Mark III, which will be available for distribution by Infotronic Systems before long. The Mark III is a completely redesigned machine. Unlike previous models, it is not restricted to the reproduction of tape recorded at 15 ips. Furthermore, if desired, the pitch of the recorded speech signal can be varied without affecting word rate. His report has not been included since it was an informal demonstration of the capabilities of the Information Rate Changer, Mark III.

The preconference workshop was presented by Dr. Willard Zemlin, Dr. Emerson Foulke, and Dr. Robert Scott. Dr. Zemlin presented a discussion of the mechanisms involved in speech production and hearing, and of acoustical energy containing speech information. Dr. Foulke explained the compression or expansion of speech by the sampling method and described the manner in which it is accomplished by electromechanical compressors of the Fairbanks type. Dr. Scott described the general procedures involved when computers are used for the compression or expansion of speech by the sampling method. The remarks of those who conducted the Workshop have not been included in this volume, since they were made extemporaneously, and since the effort to record them on tape was not entirely successful. However, no new information was presented at the Workshop. Its purpose was to provide a background for inexperienced Conference participants, and the information presented is generally available elsewhere.

CHAPTER II

AN INTRODUCTION TO SPEECH TIME COMPRESSION TECHNIQUES: THE EARLY DEVELOPMENT OF SPEECH TIME COMPRESSION CONCEPT AND TECHNOLOGY

H. Leslie Cramer*

Introduction

It should be obvious that, until it was possible to record and play back speech or sound in some manner, it was impossible to develop any sort of speech compression system. Speech time compression has only been possible and developed as the technology for mechanical and electronic acoustic recording has advanced.

There are two parallel developments that have taken place. One is the conceptual development of time compression. The second is the development of audio recording-playback systems, which, although preceding the development of the concept of time compression, will be taken up in the latter part of this paper.

The Conceptual Development of Time Compression Methods

Following are findings of some of the significant experiments that led researchers gradually into the idea of time compressing speech.

One of the earliest experimenters in this field was Harvey Fletcher (1929) of the Bell Telephone Research Laboratories. In 1929, he published his findings on accelerating speech phonographically; that is, the playing of a phonograph record at a speed faster than that at which it was recorded. Recorded speech played in this manner increases the frequency, resulting in speech which has a "Donald Duck" or "Chipmunk" effect. At moderate rates of acceleration, such speech is intelligible, especially with practice in listening to it. There have been many studies dealing with the comprehension of speech so produced. However, the remainder of this paper will be limited to the development of time compression of speech without attendent frequency distortion, or rise in pitch.

---

*Dr. H. Leslie Cramer is Senior Research Analyst, Peace Corps, 806 Connecticut Avenue, N. W., Washington, D. C. 20525.

The pattern shown in Figure 2.1 was made on a sound recording instrument called an oscillograph. This is the tracing of the vowel sound /a/, as in the word "father." A single fundamental cycle or pitch period of this voice tracing is represented by the portion between points A and B, while the portions between points B-C and C-D represent successive pitch periods. The part shown here is only a small part of a vowel sound, which may have from 20 to 50 complete cycles, depending on the pitch of a speaker's voice, his rate of speaking, and the particular vowel spoken.

Gemelli (1934) in Italy and Peterson (1939) in the United States, both experimented with the time duration of a phoneme which is necessary for it to be properly perceived. Their findings were nearly identical; that is, both discovered that only one or two complete pitch periods of a vowel sound are necessary for its perception and identification. These findings made it clear that, at least in vowel sounds, there is a high degree of redundancy in speech.

Steinberg (1936) reported that speech rates could be increased by playing records at accelerated speed without a great loss in intelligibility, at least with moderate rates of increase.

In 1940, Goldstein at Columbia University started experimenting with rate of speech to determine the comprehension of continuous discourse at gradually increasing increments of words per minute (wpm). He recorded lectures at increasing wpm rates and then presented this recorded material to students to determine how well they could understand it. The maximum rate of 325 wpm was produced by partially accelerating a phonograph record which had been recorded at 285 wpm. This was done because he was unable to find a speaker who could articulate clearly at 325 wpm. The 325 wpm presentation, according to Goldstein, was not, however, noticeably distorted. He found that his subjects had fairly good perception and understanding at this high rate. This led to the idea that our listening speed is primarily limited by the rate of speech production rather than by perceptual or cognitive structures.

Miller (1946) and Miller and Licklider (1950) experimented with an electronic switching system for interrupting speech. This process blanked out alternate portions of speech. With 50% of the speech cut out, intelligibility fell only 15%.

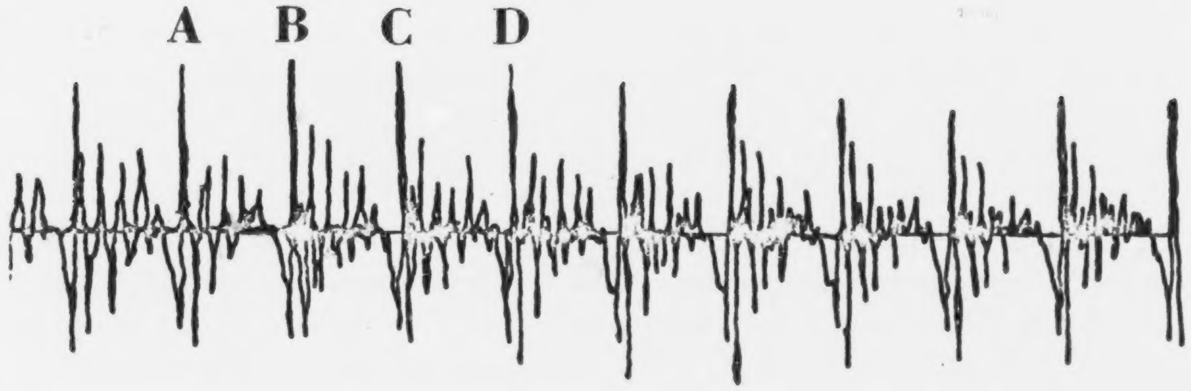


Figure 2.1. Oscillograph tracing of the /a/ sound of the word fa̲ther. (This includes only about one-fourth of the pitch periods of the /a/ sound.)

In 1948, John Black, at Ohio State University, conducted research for the Office of Naval Research (ONR). He was experimenting to determine the significance of different phonemes for word intelligibility. He simply used a razor blade to cut pieces out of a recorded tape, splicing the remaining pieces together. This was done to analyze the contribution of vowels and consonants to the intelligibility of single words.

It was Black's report on this ONR research which stimulated Garvey and Henneman (1950) to work on the "cut-splice" method. They reasoned that Black's cut and splice method could be used to eliminate part of the speech recorded on a tape, as Miller and Licklider had done in their study of electronically interrupted speech. With Black's method, however, the gaps of silence in Miller and Licklider's process would be eliminated and a saving in time would result. Their reasoning was sound, and a highly intelligible speech record was produced at speed-up ratios from 33% to 400% (1.25 to 4 times normal).

This method for time compressing speech can best be conceived by visualizing cutting alternate one-quarter inch pieces out of a recorded tape. Every other piece may be discarded, and those remaining spliced back together. Such a processed tape would make it possible to hear a half-hour lecture in 15 minutes because it is literally only half there. However, because each segment is played back at the speed at which it was recorded, there would not be the rise in pitch, or "Donald Duck" effect. Instead, the voice would sound normal in terms of pitch, and only the speed, or wpm rate, would have increased.

Figure 2.2 shows the comparison of intelligibility of "chop-splice" produced time-compressed speech with phonographically accelerated speech produced by both Garvey and Steinberg. This figure shows that the University of Virginia "chop-splice" method (Garvey) produces speech which remains above 90% intelligible at 2.5 times the input ratio. It may also be seen from this graph that the phonographic acceleration of speech by both Steinberg and Garvey does not produce speech as intelligible as the "chop-splice" method.

After finishing his thesis at the University of Virginia, Garvey was quite tired of cutting and splicing pieces of tape together. Sometime ago he stated that he was so sick and tired of recording tape and splicing tape that he hoped in his entire life he'd never see another tape splicer or reel of tape.

It is fortunate for researchers that within a couple of years of Garvey's work with the "cut-splice" method, Grant Fairbanks, W. L. Everitt, and R. P. Jaeger (1953, 1954, 1959) at the University of Illinois applied for a patent on an instrument which would automatically accomplish the same result in terms of eliminating pieces of speech.

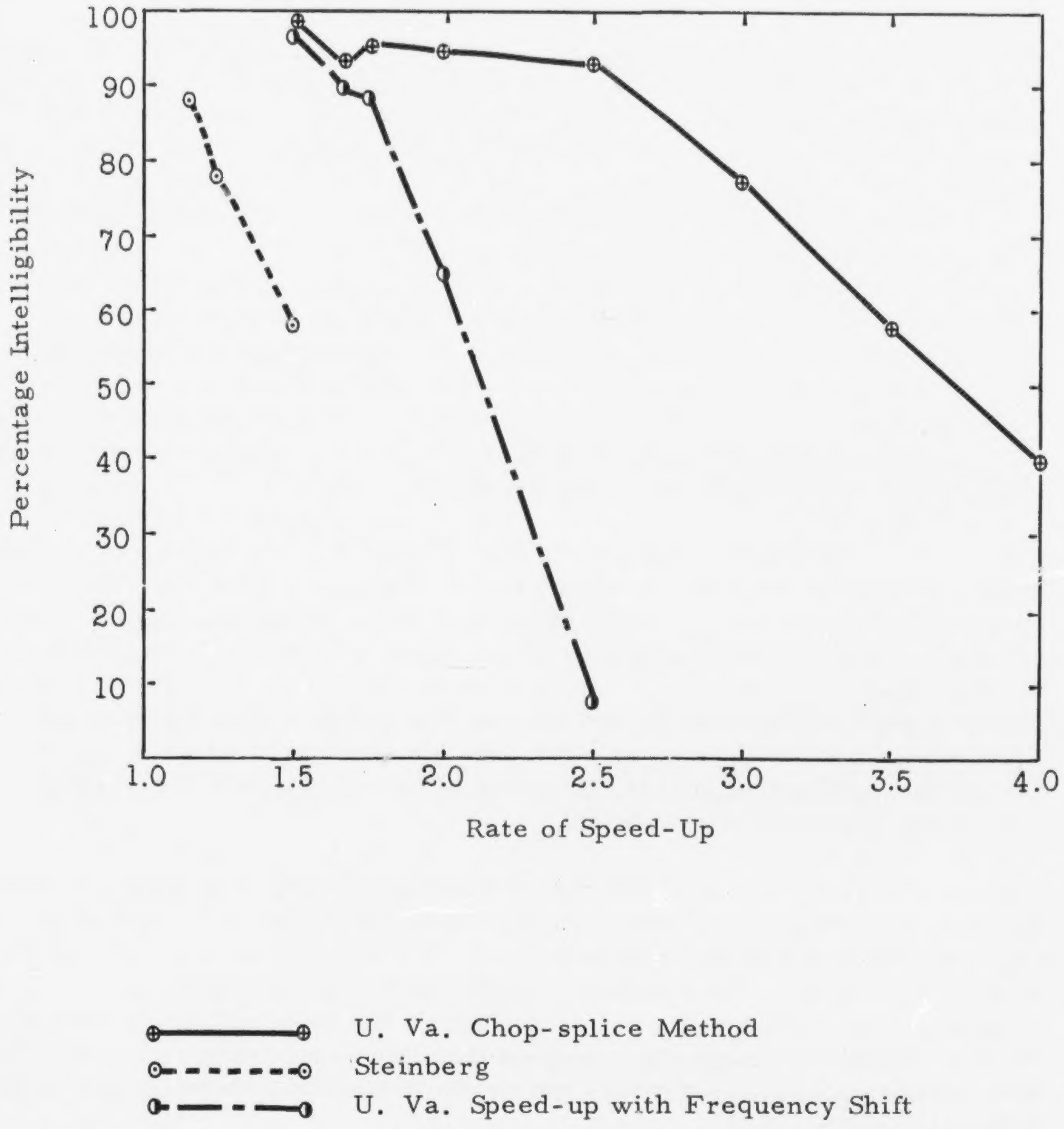


Figure 2.2. Comparison of intelligibility loss for various speed-up rates between the "chop-splice" techniques and speed-up methods involving frequency shift. (From Garvey and Henneman, 1950, p. 16, Figure 6.)

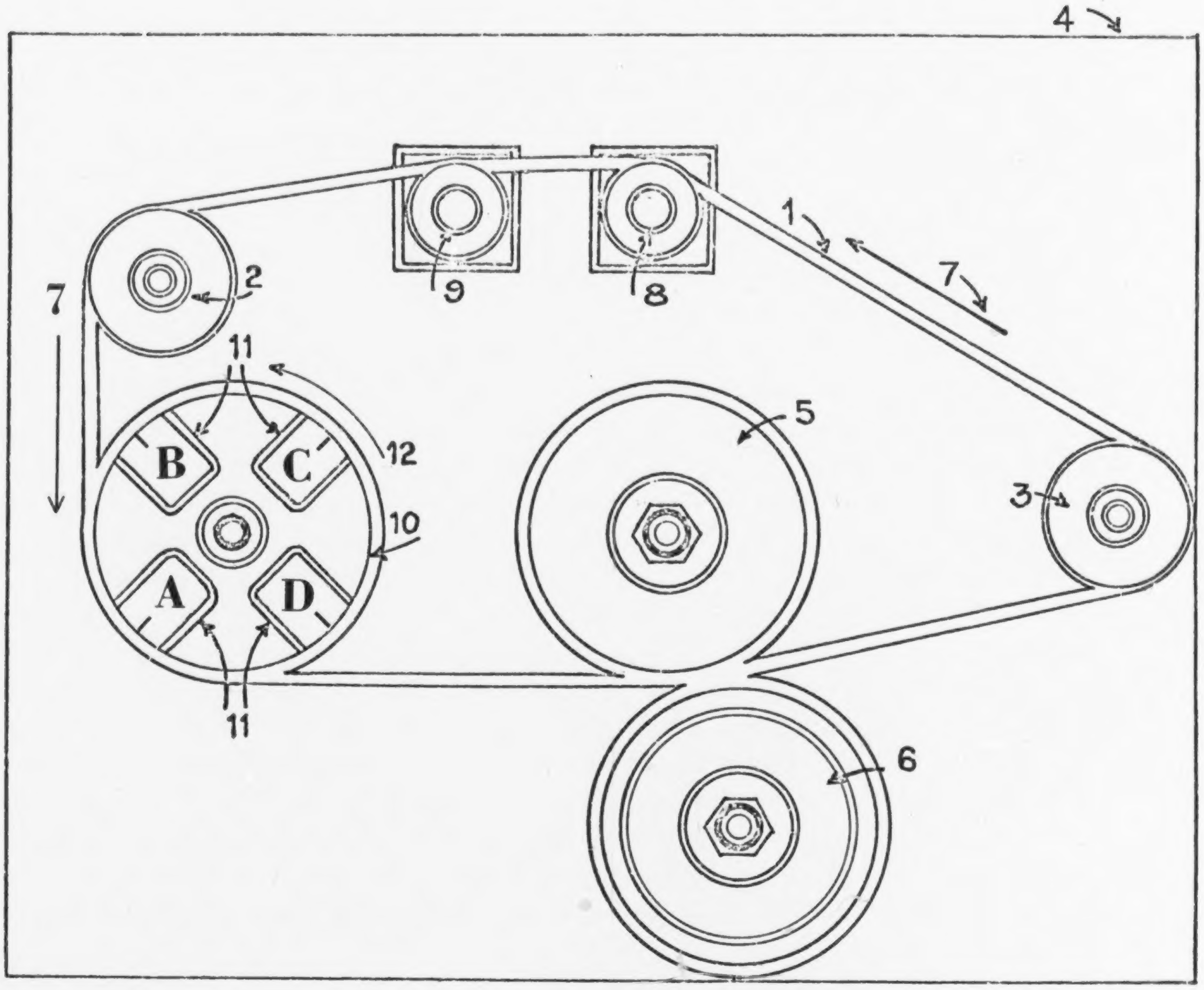
Fairbanks' method of automatically scanning a magnetically recorded tape, which reproduces a portion and eliminates another portion of each speech segment was developed by Fairbanks et al. (1953, 1954, 1959).

Referring to Figure 2.3, tape loop (1) traveling in the direction shown by arrow (7) passes over erase head (8) and recording head (9). The tape loop (1) then goes over idler (2), down around the rotating head assembly (10), between the tape drive capstan (5) and pressure roller (6), around tension adjusting wheel (3) and back to erase head (8) where it started. When the compressor is in operation, material on the tape is erased at erase head (8) in order to record cleanly at the record head (9). The recorded tape passes the rotating head assembly (10) in the direction shown by arrow (7). The tape moves faster than the rotating head assembly, so that speech recorded on the tape is picked up by any one of the four heads (A, B, C, and D) in the assembly over which it is passing. At the instant when head A leaves contact with the tape, head B contacts the tape. Everything recorded on the tape wrapped around the rotating head assembly between heads A and B will not be scanned or played back by either head A or B and therefore will be discarded. The temporal length of the unscanned material is referred to as the interval discarded (I_d), while the part played back by each head constitutes the interval sampled (I_s). These two factors can be varied with the Fairbanks equipment so that one may specify either a specified sampling or a discard interval at any given compression ratio. Of the three factors--compression ratio, discard interval, and sampling interval--two have to be fixed.*

It may be seen in retrospect that the work of Fletcher, Steinberg, and Goldstein showed that one could clearly understand speech at rates faster than speakers are capable of articulating and producing continuous discourse. Gemelli and Peterson added experimental evidence that one need hear only a small part of vowel sounds to properly identify them. Miller found that alternate portions of speech could be blanked out without a great decrease in intelligibility. Black and Garvey both eliminated pieces of the speech record without leaving blank spaces so that the wpm rate was increased without a pitch rise or great loss in intelligibility. The final synthesis of these findings was their embodiment in the speech compressor invented by Fairbanks et al. (1959).

---

*A more complete explanation of Fairbanks' Compressor, complete with operating formula and peripheral equipment adjustments, is available in Cramer (1968), pp. 40-51 and pp. 191-203.



1 - Tape Loop
2 - Idler Wheel
3 - Tension Adjusting Wheel
4 - Mounting Plate
5 - Capstan
6 - Pressure Roller
7 - Direction of Tape Loop Travel

8 - Erase Head
9 - Record Head
10 - Rotating Head Assembly
11 - Playback Heads
12 - Direction of Rotating Head Assembly when Compressing

Figure 2.3. Detail drawing of Fairbanks' compressor. (From Fairbanks, Everitt, and Jaeger, 1959.)

The Development of the Technology for Time Compressing Speech

The second major development referred to at the beginning of this paper relates to the technology for electromechanical recording and playback of auditory signals. The treatment of this area must necessarily be restricted to that bearing directly on methods of either recording or playback that scan or sample an original auditory input, or otherwise translate frequencies up or down the scale. The writer believes that the coverage of this topic is complete, but will be most interested in references to any other devices on which patents are held that are not reported here.

In the following brief review of patents, the dates given in the text will be the original filing dates, while those in the references represent the actual date a patent was issued. This seems necessary in view of the fact that it is the date of conception of the idea that is important, and in many cases there was a substantial delay in the awarding of the patent. However, the interested reader needs the date of issue in order to retrieve information on the patents.

The earliest record of a speech scanning system is a U. S. patent filed by N. R. French and M. K. Zinn (1928) in December 1924. This system proposes to rotate a microphone around a sound pipe or speaking tube bent in a circle with a slot around its edge (see Figure 2.4).

This patent, it turns out on analysis, would not work with air as the sound carrying medium, since the tube would have to be 15 feet in circumference and scanned at 32,000 rpm in order to accomplish 50% compression. This patent therefore really only represents the concept of scanning without the reduction to practice normally required in a patent.

In 1930, Berthold Freund (1935) applied for a patent on a device used for scanning motion picture film sound recordings which could be used to vary the length of sound records. This was developed for synchronizing sound to the film track and for shortening the speech record to match a speeded up portion of film, without having a pitch rise. This appears to be the first apparatus capable of actually time compressing speech, although no claim for use other than to match film records was made (see Figure 2.5).

In 1935, Homer W. Dudley (1938) applied for a patent on a signalling system which sampled every other pitch period of speech and transmitted it to a distant point. At that distant point, each pitch period was repeated once to reestablish a wave form similar to the original. The patent made no claim for saving the listener time, as it was only to save time on transmission that it was developed. This same apparatus could, without repeating each signal on the output end, compress speech.

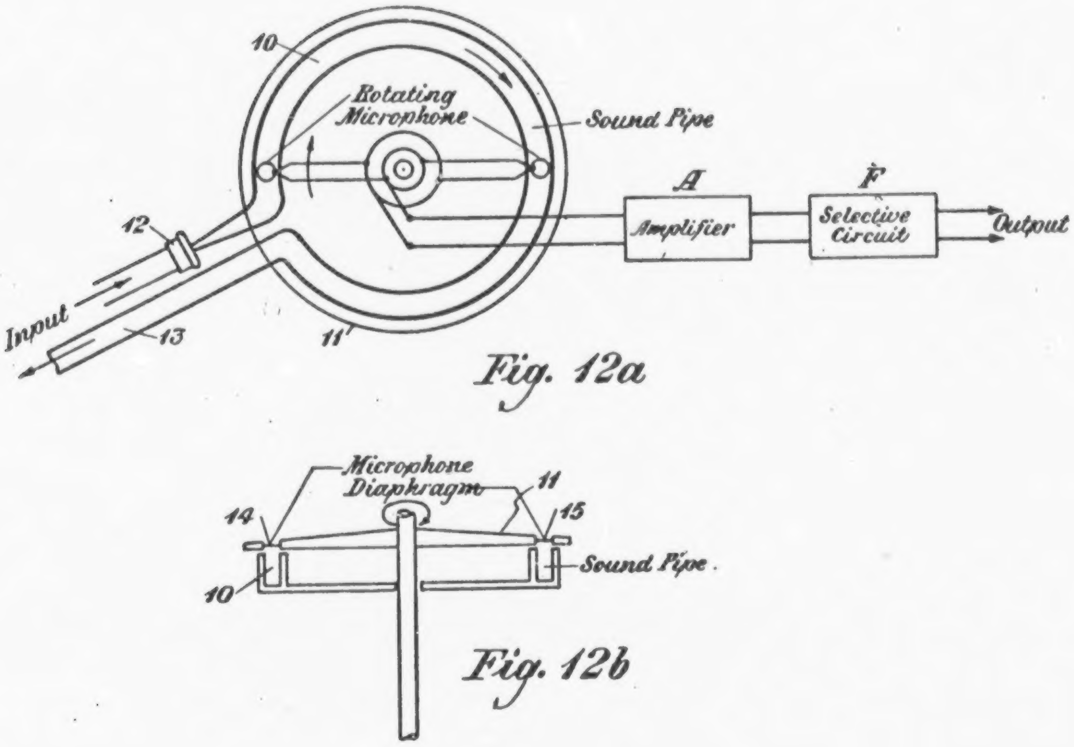

10
Rotating Microphone
Sound Pipe
A
F
Amplifier
Selective Circuit
Output
12
Input
13
11'
Fig. 12a
Microphone Diaphragm
11
14
15
Sound Pipe
10
Fig. 12b

Figure 2.4. Detail from French and Zinn (1928), showing their Figure 12a and Figure 12b illustrating a rotating microphone sound tube scanning system.

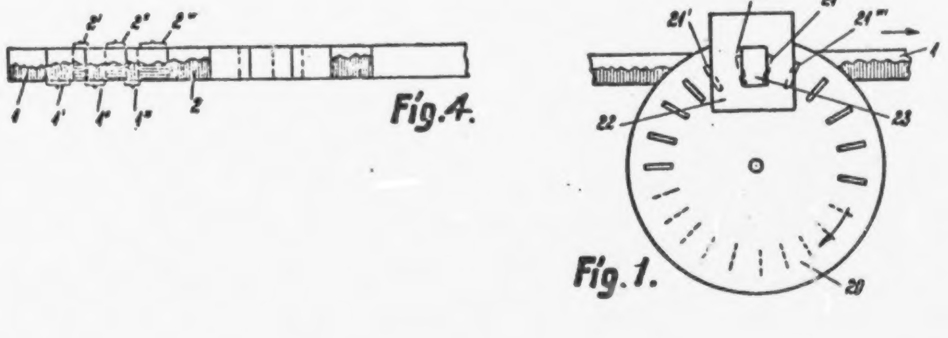

2'
2''
2'''
1
1'
1''
1'''
2
Fig.4.
21'
21''
21'''
21''''
1
22
23
20
Fig.1.

Figure 2.5. Detail from Freund (1935), showing his Figure 1 and Figure 4 illustrating system for scanning motion picture film sound track.

In 1936, R. L. Miller (1939) applied for a patent on a signalling system which used frequency division for speech bandwidth reduction. This patent anticipated the harmonic compressor as worked out recently by the American Foundation for the Blind.

In 1936, Leonid Gabrilovitch (1939) developed a system for scanning a steel wire recording with rotating heads. This was similar to Dudley's in that it was designed to reduce frequencies for transmission. At the transmitting end of a line, every other segment was divided before transmission, then at the receiving end it was multiplied in frequency and repeated once before the next segment arrived (see Figure 2.6).

In 1938, Eduard Schüller (1942) patented a similar device in Germany which was used for playing back magnetic recordings in less time or in longer time than that in which they were actually recorded. In his patent he states:

> "If the sound head is rotated in the same direction as that of the travel of the record strip, . . . an acoustical time compressing is obtained and the reproduced signal has its original frequency but is read off in less time than was required for the recording."

This is the first clear reference to time compressing speech with the method Fairbanks later developed, apparently independently.

Figure 2.7 shows Figures 1 and 2 from Schüller's patent and greatly resembles others.

In 1944, Gabor (1949) applied for a patent on a device using microscope lenses in a ring to scan the sound track of a motion picture film. See Figure 2.8 for Gabor's diagram of this process.

In 1947, Gabor (1950) developed many ingenious ways of both scanning and blending adjacent samples of the speech record. See Figure 2.9 for diagrams of some of these. These figures display systems of scanning a track photographically, and electronically. Gabor's Figure 11 shows how lenses are formed by discharging a spark synchronized to voice pitch periods, through water. The bubbles of gas so produced are then circulated by a pump past the sound record to be scanned. A real Rube Goldberg device!

In 1950, Vilbig (1950, 1952, 1967) described a string filter device which is a physical analog of the electronic harmonic compressor. It could be used only to compress by a factor of two to one. Figure 2.10 shows a picture of this complex system.

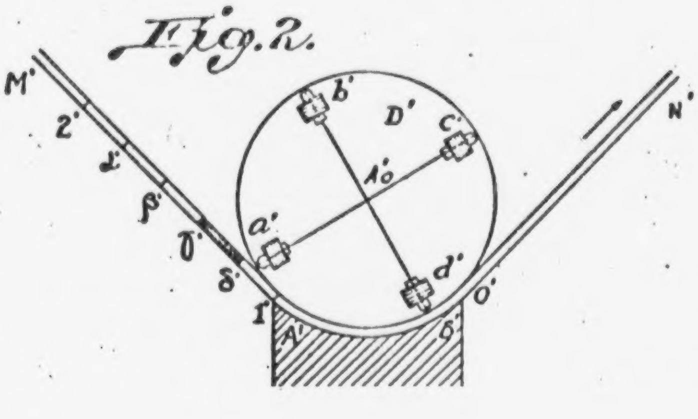


Fig. 2.
M'
N'
D'
b'
c'
a'
d'
A'
B'
O'

Figure 2.6. From Gabrilovitch (1939), showing his Figure 2 illustrating the rotating scanning heads.

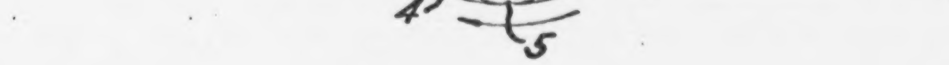


Fig. 1.
a
a
1
2
3
4
5
6

Figure 2.7. From Schuller (1942), showing his Figure 1 illustrating his rotating magnetic pickup heads.

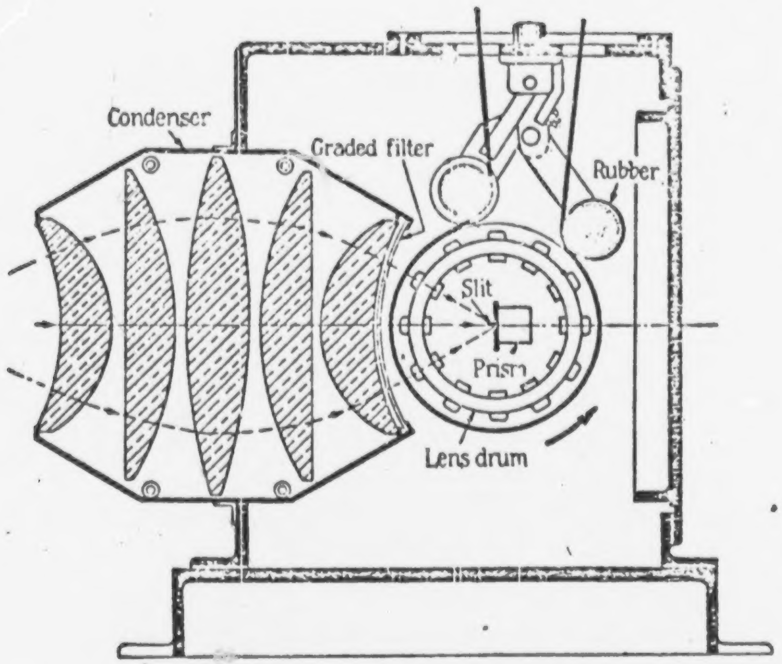

Condenser
Graded filter
Rubber
Slit
Prism
Lens drum

Figure 2. 8. From Gabor (1949), Figure 1 illustrating his lens drum used to scan motion picture film tracks.

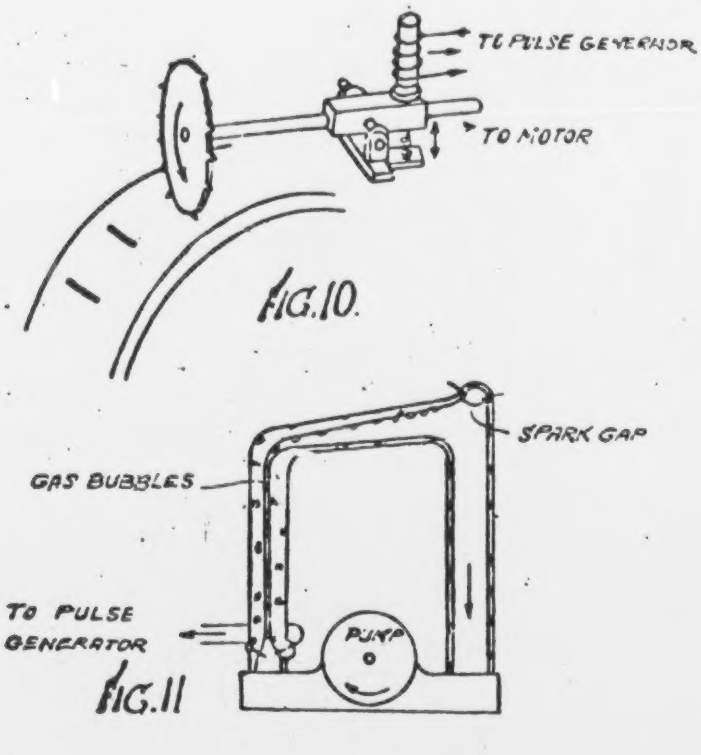

TO PULSE GENERATOR
TO MOTOR
FIG.10.
SPARK GAP
GAS BUBBLES
TO PULSE GENERATOR
PUMP
FIG.11

Figure 2. 9. From Gabor (1950), Figure 10 and Figure 11 illustrating his apparatus for scanning a motion picture film track in sychrony with voice pitch periods.

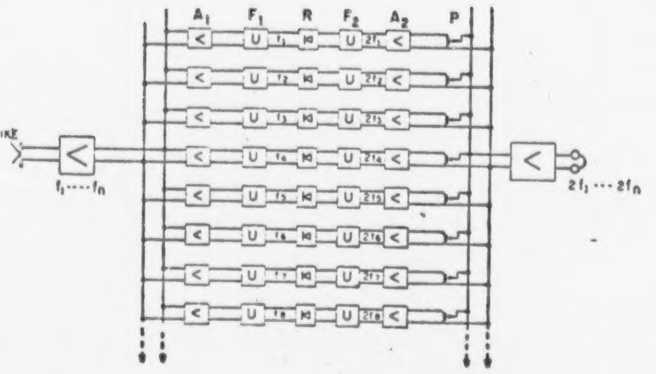

A1
F1
R
F2
A2
P
MIKE
f1....fn
2f1...2fn

FIG. 1. Schematic diagram of the basic circuitry of a "distortion-free" frequency doubler.

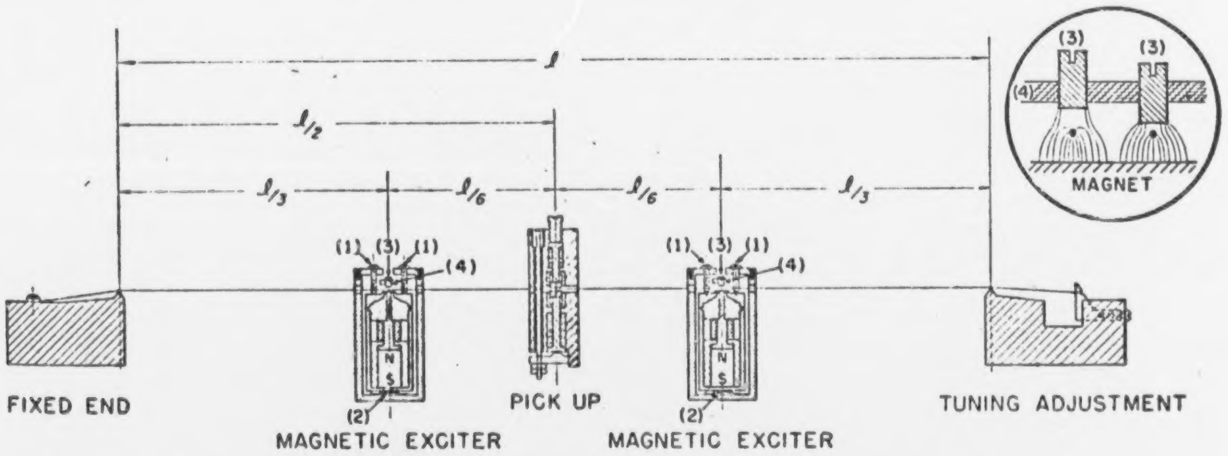

l
l/2
l/3
l/6
l/6
l/3
(1)
(3)
(4)
(2)
MAGNET
FIXED END
MAGNETIC EXCITER
PICK UP
MAGNETIC EXCITER
TUNING ADJUSTMENT

FIG. 2. Construction of the exciting coils and the pick-up in a cross-sectional view.

Figure 2.10. From Vilbig (1950), Figure 1 and Figure 2 illustrating his string filter analog of the harmonic compressor.

In the fall of 1952, Grant Fairbanks et al. (1959) applied for their patent on the compressor system* developed at the Speech Research Laboratories at the University of Illinois. This system uses the rotating head assembly shown in Figure 2.3.

Anton Springer (1961a, 1961b; 1962a, 1962b, 1962c; 1963) filed a series of patents on improvements on the rotating heads and driving mechanisms starting in 1956. These have been incorporated into the Eltro Tempo Regulator manufactured in Germany.** These machines have the advantage of a continuously adjustable compression rate up to 1.7 times normal wpm rate, but a disadvantage in terms of the long discard interval of 40 milliseconds.

Schimmel and Clay (1963) filed for additional improvements on rotating heads. This was mainly an air suspension system to reduce tape and head wear. Gabor (1965) patented a multihead system with provision for synchronizing the sampling to the occurence of pitch periods in the speech record being processed.

Robert J. Wenzel (1962) working at Massachusetts Institute of Technology with John Dupress, developed a jitter action time compression device using the ignition timing cam from an automobile as the basic driving device. This did not work too well due to mechanical vibration but may well deserve renewed effort as it would be an inexpensive system to produce.

Jay Harold Ball (1961) developed the first known computer program for compressing speech. His work was followed by Scott (1965), H. L. Cramer and R. P. Talambiras (1970), and S. U. Qureshi and Y. J. Kingma (1970).

There are reportedly two solid state systems under development using essentially a long taped delay line for slowing speech and thereby reducing frequencies below normal.

---

*This device was first commercially available as the Vari-Vox machine, manufactured by Kay Electronics, Inc. in New Jersey. It is now commercially available in improved form from Discerned Sound, Inc., North Hollywood, California. Samples of the speech produced on this type of machine are available from the Center for Rate-Controlled Recordings, University of Louisville, Louisville, Kentucky. This Center also has facilities for processing tapes at any specified amount of compression at a nominal fee.

**Available from Gotham Audio, New York, New York.

To summarize, in terms of available systems today, there are three different approaches. First, in terms of both discovery and amount of usage today, is the rotating head assembly system of Fairbanks et al. Secondly, we have several computer programs, somewhat costly and not generally available. Thirdly, we have the Harmonic Compressor developed by the American Foundation for the Blind and now available at the Perceptual Alternatives Laboratory at the University of Louisville, Louisville, Kentucky.

CHAPTER III

EFFECT OF RATE OF COMPRESSION AND MODE OF PRESENTATION ON THE COMPREHENSION OF A RECORDED COMMUNICATION TO JUNIOR COLLEGE STUDENTS OF VARYING APTITUDES

Clement Cordell Parker*

A problem common to most educational institutions is to find better techniques to send information across media with speed and reliability. The problem is aggravated within junior colleges because of the increased heterogeneity of its student population.

A number of studies have been made to determine the relationship of rate of presentation with degree of comprehension. Harwood (1955) discovered an insignificant loss as word rate was increased. Fairbanks, Guttman, and Miron (1957c) found little difference in the comprehension of messages presented at 141, 201, and 282 words per minute (wpm). The results of these and other studies seem to indicate that while there is a loss in comprehension with an increase in rate of presentation, the loss is insignificant up to about 280 wpm.

Sticht (1968) trichotomized 135 Army inductees into three mental aptitude categories--low, medium, and high--according to their Air Force Qualification Test scores. He found that increasing the speech rate had a greater disrupting effect on test performance of the higher aptitude subjects than those of low aptitude.

Travers (1964) reports that he and Jester presented reading passages through hearing alone, vision alone, and both hearing and vision. They found that at the slower speeds no advantage was found for the audiovisual presentation, but at higher speeds the audiovisual channel proved to be superior. Loper (1966) measured comprehension and retention using two modes: aural and visually augmented aural where televised pictorials were used to supplement the aural message. He concluded that visual augmentation does not provide much assistance to an aural presentation.

---

*Mr. Parker is Chairman of the Department of Speech & Drama on the Northeast Campus of the Tarrant County Junior College District in Fort Worth, Texas. He is a candidate for the Doctor of Education degree at North Texas State University in Denton, Texas.

Junior college students score lower on aptitude tests than those students in four year colleges. The research (Cross, 1968) is national in scope, unanimous in findings, and is based on a staggering array of accepted measures of academic aptitude.

This study was conducted with the hope that an efficient method for processing information for junior college students could be discovered, thereby increasing their learning and success potential.

Statement of the Problem

The problem of this study was to find a more efficient way to store and transmit recorded information, thereby increasing the efficiency of programmed learning centers and reducing the time required for utilization. More specifically, the problem was to determine the rate of compression and mode of presentation having the most favorable impact on the comprehension of a recorded communication to junior college students of varying aptitudes.

Subproblems included the following: (1) determination to what degree rate of compression could be increased without significant loss in comprehension, (2) determination to what degree rate of comprehension could be increased with the simultaneous presentation of compressed speech and the printed page, and (3) determination of the effects of rate of compression and mode of presentation to students representing all levels of aptitude, low levels of aptitude, and high levels of aptitude.

Definition of Terms

1. Compressed speech--oral, tape-recorded communication in which brief segments of the message have been deleted without significant distortion in vocal pitch or quality. (1a) Zero compression, normal speaking rate; (1b) one-third compression, compressed speech requiring two-thirds of the original time for presentation; (1c) one-half compression, compressed speech requiring half of original time for presentation.

2. Audio-ocular--the addition of the printed page to match an aural message in order to add the factor of sight to a factual presentation.

3. Test of comprehension--the correct number of responses to the comprehension test within Form B of the 1960 edition of the Nelson-Denny Reading Test.

4. Test of aptitude--the correct number of responses to the Verbal Comprehension section of the Guilford-Zimmerman Aptitude Survey. (4a) All-levels group, included all students participating in experiment minus those taken from the initial sample because of absence during one of the tests or failure to hear all of the selections; (4b) high-level group, those students who, within their treatment condition, scored at or above the sixty-seventh percentile on test of aptitude; (4c) low-level group, those students who, within their treatment condition, scored at or below the thirty-third percentile on test of aptitude.

Procedure

The eight selections within the test of comprehension were recorded by a professional speaker, and compressed to one-third and one-half degrees, by the Center for Rate-Controlled Recordings at the University of Louisville. Compression was achieved through the use of a Fairbanks type compressor. Instructions and a 2-minute practice selection were programmed into the tapes.

Subjects were 429 students enrolled in the Freshman composition classes during the fall semester of the 1969-70 academic year on the Northeast Campus of the Tarrant County Junior College District in Fort Worth, Texas. Eighteen of the 22 available day sections were selected at random, and a table of random numbers utilized to populate the six experimental groups with three sections in each group (about 75 students for each of the six experimental groups).

The test of comprehension was administered during the first week of classes in the Language Laboratory within the Programmed Learning Center. Students were free to select any one of the 30 available carrels, and each carrel was equipped with padded earphones which could be adjusted for comfortable listening. Each of the output units was locked into the channel selected for the experiment. A copy of the test of comprehension was available to all audio-ocular groups, and included the printed copy of each of the recorded messages. The aural-only groups received only the test questions. Each carrel was supplied with pencil, answer sheet, and short questionnaire.

When all Ss were seated, they were asked to place their earphones on their heads. Programmed instructions began immediately thereafter with an admonition to adjust the earphones for comfortable listening and to confirm ability to hear. Students then heard the 2-minute introductory message and the eight selections of the test of comprehension. Students were allowed 15 seconds per question to answer each of the 36 multiple-choice items. When the last test was finished, students were asked to fill out a brief questionnaire and were thanked for their participation.

Subjects were also given the test of aptitude during the first week of classes. All of the tests were administered by the same person in comparable classrooms. All tests were hand-scored and the results recorded on keypunch worksheets.

Treatment of Data

Three 3 x 2 classifications of data were created. The 3 x 2 schema represented two modes of presentation (aural-only and audio-ocular) and three degrees of compression (zero, one-third, and one-half). The first 3 x 2 classification represented the all-levels group, the second the high-level group, and the third the low-level group. Two-way analysis of variance yielded the following results:

TABLE 3.1

TWO-WAY ANALYSIS OF VARIANCE FOR TEST OF COMPREHENSION ALL-LEVELS GROUP

Source of Variation	SS	df	MS	F
Mode of Presentation	1,428.35	1	1,428.35	54.73*
Rate of Compression	1,101.43	2	550.71	21.10*
Interaction	349.97	2	174.99	6.70*
Within	11,040.34	423	26.10	

*p < 0.05

TABLE 3.2

TWO-WAY ANALYSIS OF VARIANCE FOR TEST OF COMPREHENSION HIGH-LEVEL GROUP

Source of Variation	SS	df	MS	F
Mode of Presentation	671.79	1	671.79	35.16*
Rate of Compression	291.48	2	145.74	7.63*
Interaction	253.63	2	126.81	6.64*
Within	2,617.42	137	19.10	

*p < 0.05

TABLE 3.3

TWO-WAY ANALYSIS OF VARIANCE FOR TEST OF COMPREHENSION LOW-LEVEL GROUP

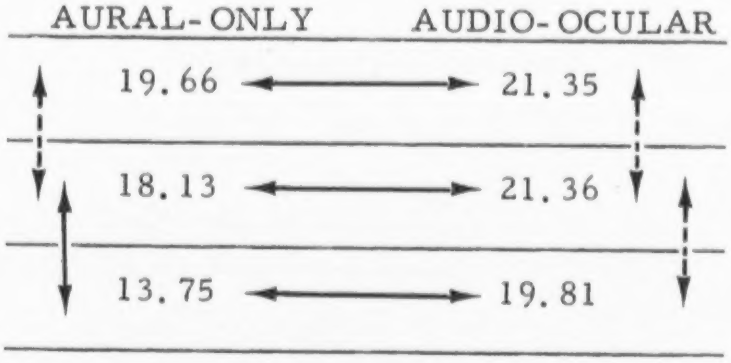
Source of Variation	SS	df	MS	F
Mode of Presentation	204.43	1	204.43	10.31*
Rate of Compression	424.72	2	212.36	10.71*
Interaction	43.67	2	21.83	1.10
Within	2,716.85	137	19.83	

*$p < 0.05$

Since the results from the analysis of variance permitted rejection of all null hypotheses of no difference due to rate of compression or mode of presentation at different aptitude levels, *t* tests were run for comparison of certain means with the following results:

TABLE 3.4

MEAN COMPREHENSION SCORES ALL-LEVELS

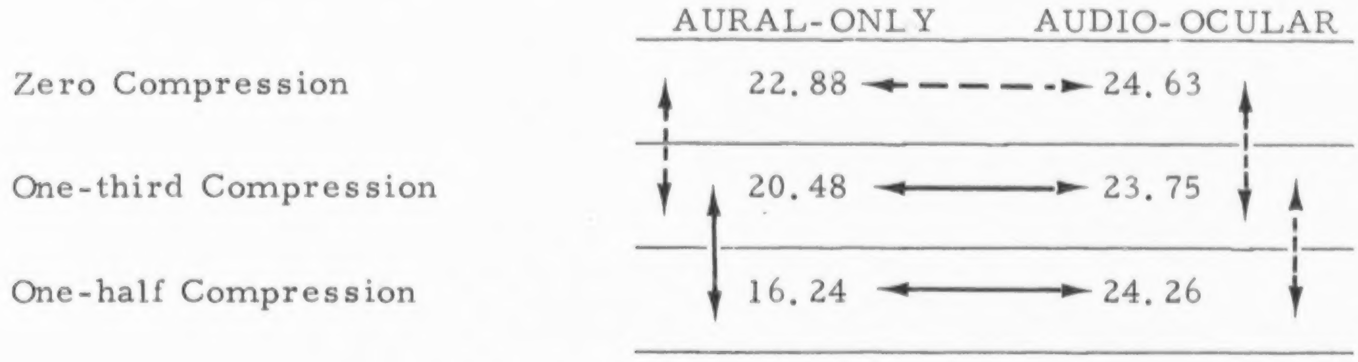


	AURAL-ONLY	AUDIO-OCULAR
Zero Compression	19.66	21.35
One-third Compression	18.13	21.36
One-half Compression	13.75	19.81

Key: ←- - - - -→ No significant difference between means
←———→ Significant difference at .05 level or better.

TABLE 3.5

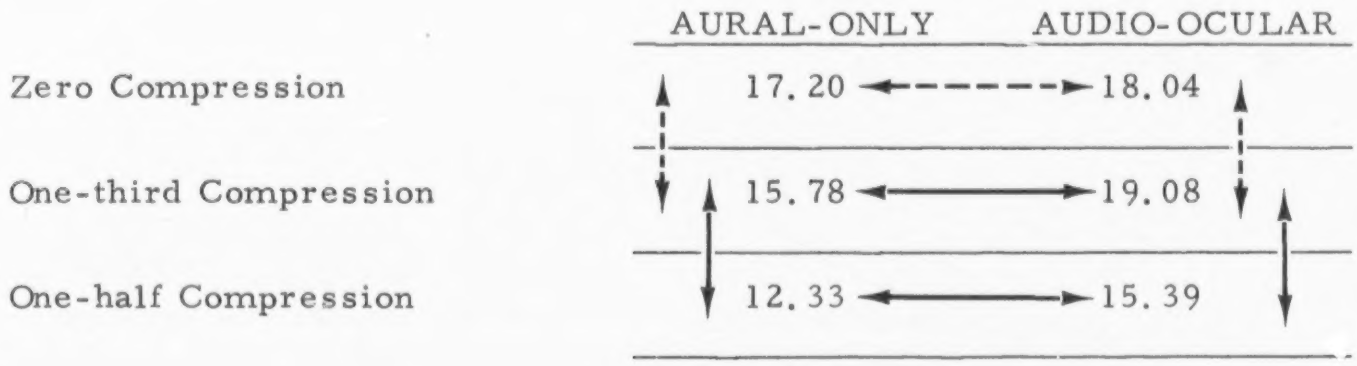
MEAN COMPREHENSION SCORES HIGH-LEVEL



	AURAL-ONLY		AUDIO-OCULAR
Zero Compression	22.88	←- - - - -→	24.63
One-third Compression	20.48	←————→	23.75
One-half Compression	16.24	←————→	24.26

TABLE 3.6

MEAN COMPREHENSION SCORES LOW-LEVEL



	AURAL-ONLY		AUDIO-OCULAR
Zero Compression	17.20	←- - - - - -→	18.04
One-third Compression	15.78	←————→	19.08
One-half Compression	12.33	←————→	15.39

Key:

←- - - - - -→ No significant difference between means.

←————→ Significant difference at .05 level or better.

TABLE 3.7

_t_ TESTS FOR COMPREHENSION SCORES ALL-LEVELS GROUP

Run	Mean	N	Mean	N	df	t
1	19.66	76	21.35	80	154	-2.07*
2	18.13	68	21.36	73	139	-3.74*
3	13.75	63	19.81	69	130	-6.81*
4	19.66	76	18.13	68	144	1.79
5	18.13	68	13.75	63	129	4.91*
6	21.35	80	21.36	73	151	- .01
7	21.36	73	19.81	69	140	1.80

$^*p < 0.05$

TABLE 3.8

t TESTS FOR COMPREHENSION SCORES HIGH-LEVEL GROUP

Run	Mean	N	Mean	N	df	t
1	22.88	25	24.63	27	50	-1.44
2	20.48	23	23.75	24	45	-2.56*
3	16.24	21	24.26	23	42	-6.08*
4	22.88	25	20.48	23	46	1.90
5	20.48	23	16.24	21	42	3.21*
6	24.63	27	23.75	24	49	.72
7	23.75	24	24.26	23	45	- .40

*$p < 0.05$

TABLE 3.9

t TESTS FOR COMPREHENSION SCORES LOW-LEVEL GROUP

Run	Mean	N	Mean	N	df	t
1	17.20	25	18.04	27	50	- .68
2	15.78	23	19.08	24	45	-2.54*
3	12.33	21	15.39	23	42	-2.27*
4	17.20	25	15.78	23	46	1.10
5	15.78	23	12.33	21	42	2.57*
6	18.04	27	19.08	24	49	- .83
7	19.08	24	15.39	23	45	2.84*

*$p < 0.05$

Discussion

The simultaneous presentation of the printed page to match an aural presentation resulted in significantly better comprehension for all aptitude levels hearing compressed speech. It was not, however, superior for the high and low aptitude level groups hearing normal rate recordings. Hence, it may be concluded that the printed page provides assistance in comprehension when the speaking rate is increased above the normal rate.

None of the aptitude levels experienced significant losses in comprehension when messages were speeded to one-third compression. This illustrates the suitability and efficiency of compressed speech for a junior college population. Furthermore, except in low-aptitude groups, the speed may be increased to one-half compression without significant loss in comprehension, provided the printed page is supplied to match the aural message. Comprehension was significantly decreased when the aural-only messages were speeded to one-half compression. A speed of one-half compression may be too great to result in acceptable comprehension for aural-only groups.

CHAPTER IV

PERTURBATIONS OF SEX JUDGMENTS WITH TIME-COMPRESSED AND FREQUENCY-DIVIDED SPEECH SIGNALS

Daniel S. Beasley and Willard R. Zemlin*

If time-compressed and frequency-divided speech is to be used in educational and clinical settings, the equivocal results of several studies of subjective perceptual interpretation of the processed speech signal should be investigated.

Time-Compressed Speech

Daniloff, Shriner, and Zemlin (1968a) observed female speakers to be rated as more intelligible than male speakers when they spoke eight vowels in an h-d context which were time-compressed using the Fairbanks sampling method. However, Zemlin, Daniloff, and Shriner (1968) also showed that listeners rated female time-compressed speech as more difficult to listen to than male time-compressed speech. In addition, the same judges preferred 30% time-compressed speech over 40% and 50%, although the Daniloff et al. (1968a) study showed that intelligibility was high up to compression rates of 70%. It appears phonemic quality, as reflected by vowel intelligibility, may remain more stable at higher compression ratios than phonetic quality. Foulke (1966c) distributed recordings of time-compressed speech and questionnaires to blind Ss in several geographical areas. Although the majority of the respondents found the female easier to understand than the male (55% versus 45%), a larger majority preferred to listen to the male (65% versus 32%). These results suggest that speaker preference criteria of auditory Os may play an equal if not greater role in the utilitarian consideration of such speech. Evidence has been provided that phonemic quality is based on a relative vowel hypothesis (Daniloff et al., 1968a; Potter & Steinberg, 1950), whereas phonetic quality may be based on a modified fixed vowel hypothesis as suggested by Slawson (1968). Phonemic quality of female speech would be maintained longer than male speech due to the inherent redundancy of female speech, but phonetic quality would decline earlier for female speech because more of the characteristic pitch periods (determining fundamental frequency) of the female, contra the male, are discarded in the sampling technique. Listener preference may be partially determined by this phonetic quality. Good phonemic quality may not overcome listener's dislike of listening to the material in a prolonged listening task. It is then necessary to study preference values of the listeners using time-compressed female speech in order to establish possible reasons the male is preferred over the female. Such knowledge would perhaps lead to methods of overcoming these attitudes, thereby permitting, in the educational process, full advantage to be taken of the high intelligibility of female time-compressed speech.

---

*Dr. Daniel S. Beasley is an Assistant Professor in the Department of Audiology and Speech Sciences at Michigan State University, East Lansing, Michigan 48823. Dr. Willard R. Zemlin is Director of the Speech and Hearing Research Laboratory at the University of Illinois, Champaign, Illinois 61820.

Frequency-Divided and Frequency-Divided Time-Restored Speech

Daniloff et al. (1968a), in their vowel study, showed female frequency-divided and frequency-divided time-restored speech had better phonemic quality than male speech, as did Klumpp and Webster (1961) using a slow playback frequency-divided method. However, neither looked at preference values for frequency-divided and frequency-divided time-restored speech. Bennett and Byers (1967) investigated the use of frequency-divided speech, using a slow playback method, on a geriatric population. Their Ss preferred the male speech. Thus, sex of the speaker may yield differential results for phonemic and phonetic quality in studies involving frequency-divided speech. Based on the relative vowel hypothesis, phonemic quality of female speech may remain higher than the male's, since the female's lower formants, especially F2 (Thomas, 1968), unlike the male's, are not shifted out of the normal experiential bandpass under frequency-divided and frequency-divided time-restored conditions (Daniloff et al., 1968a; Tiffany & Bennett, 1961). But the formant shifting does effect phonetic quality, which is based on fixed values. In a prolonged listening task, phonetic quality must be considered. The reason for the above conflicting results may be that the more intelligible frequency-divided and frequency-divided time-restored female speech, when shifted toward the frequency domain of the male, begins to sound effeminate, a cultural taboo in our society, or at least it used to be, and members of the society, as listeners, may not prefer to listen to it.

The purpose of this study is to investigate the ratings of masculine-feminine continuum poles of a male and female speaker whose speech has been time-compressed, frequency-divided, and frequency-divided time-restored. The masculine-feminine data will be compared to values obtained on other scales in similar studies.

Method of Investigation

Experimental Materials

In order to adequately compare the phonemic analysis of Daniloff et al. (1968a) to the phonetic analysis of this study, the stimuli consisted of 11 h-d context embedded vowels, spoken by a male (fo = 104 Hz) and a female (fo = 198 Hz) at conversational pitch and effort level. The vowels were processed through five conditions (20% through 60% in 10% steps) of time-compressed and frequency-divided and frequency-divided time-restored speech. Thus, there were 32 experimental sets of vowels: 2 normal (male and female); 10 time-compressed (5 males, 5 females); 10 frequency-divided (5 males, 5 females); and 10 frequency-divided time-restored (5 males, 5 females).

The 32 sets of vowels were randomized. All Os heard the same randomized experimental tape. Approximately 2 seconds of silent interval was provided between items in each set. Each set of words took about 25 seconds playback time.

Subjects

Listeners consisted of 14 male and female college students in a controlled listening environment.

Experimental Procedures

Semantic differential type scales (Osgood, Suci, & Tannebaum, 1957) were used to assess phonetic quality. These attempt to elicit behavior to alternatives which are representative of the various meanings over which a concept (in this case, speech sample) may vary on a 7-point scale of polar opposites to indicate direction and intensity of response. Seven such semantic differential scales, chosen according to the Osgood et al. (1957) criteria of relevance (of the scales to the concepts being judged) and linearity of polar opposites (e. g., rugged-delicate may both be favorable under certain circumstances), were used to elicit qualitative judgment of the 32 sets of speech signals from the listeners. These seven scales were: Fast-Slow, High-Low, Masculine-Feminine, Like-Dislike, Harmonious-Dissonant, Loud-Soft, Pleasant-Unpleasant.

The O's task was to rate each of the 32 sets of each of the seven scales. Observer heard a set, then was allotted 1 minute to respond to the 11 items in the set. A 1 minute response interval was used to allow the O adequate response time on more difficult sets. Further, the long interval aided in the forgetting of prior sets, thus minimizing the tendency of O to compare subsequent sets to prior sets. Prior to the beginning of a set, three 1 kHz beeps were sounded as a warning to "get ready." A single beep sounded at the end of a set indicating the 1 minute rating period had begun.

The response sheet consisted of three scale-position randomizations (R1, R2, and R3). These three randomized sheets were randomly distributed in booklets of 41 each for each O. Finally, the poles on the continua for R1, R2, and R3 were randomly positioned, so that one end (left or right) of the continua was not always positive and/ or negative.

All Os received standardized instructions (see Appendix A).

Phase II of the study was similar to Phase I, except an Intelligible-Unintelligible scale was added to the rating sheets. A different male and female speaker was used, thus bringing the total number of speakers to four: two males and two females. Also, Phase II eliminated ratings of time-compressed speech. Finally, 15 different listeners were used in Phase II.

Results--Phase I

Reliability of Ratings

An intraclass correlation coefficient (McNemar, 1962) for the masculine-feminine continuum was computed for the total group. The r_{tt} was found to be .99.

$\overline{M}$ Values of Ratings by Conditions

Table 4.1 lists the $\overline{M}$ scale values by condition, by set, and by sex of speaker. Figures 4.1, 4.2 and 4.3 illustrate the values of Table 4.1 graphically.

As can be seen from Figure 4.1, the male and female speakers are consistently (r_{tt} = .99) rated as per their respective sex under time-compressed speech. The high r_{tt} suggests that the variations in the M ratings by sets for time-compressed speech are systematic. There appears to be a trend toward middle scale values for both speakers, the male showing the trend sooner, but the female showing the trend more consistently, especially at higher time-compressed speech ratios.

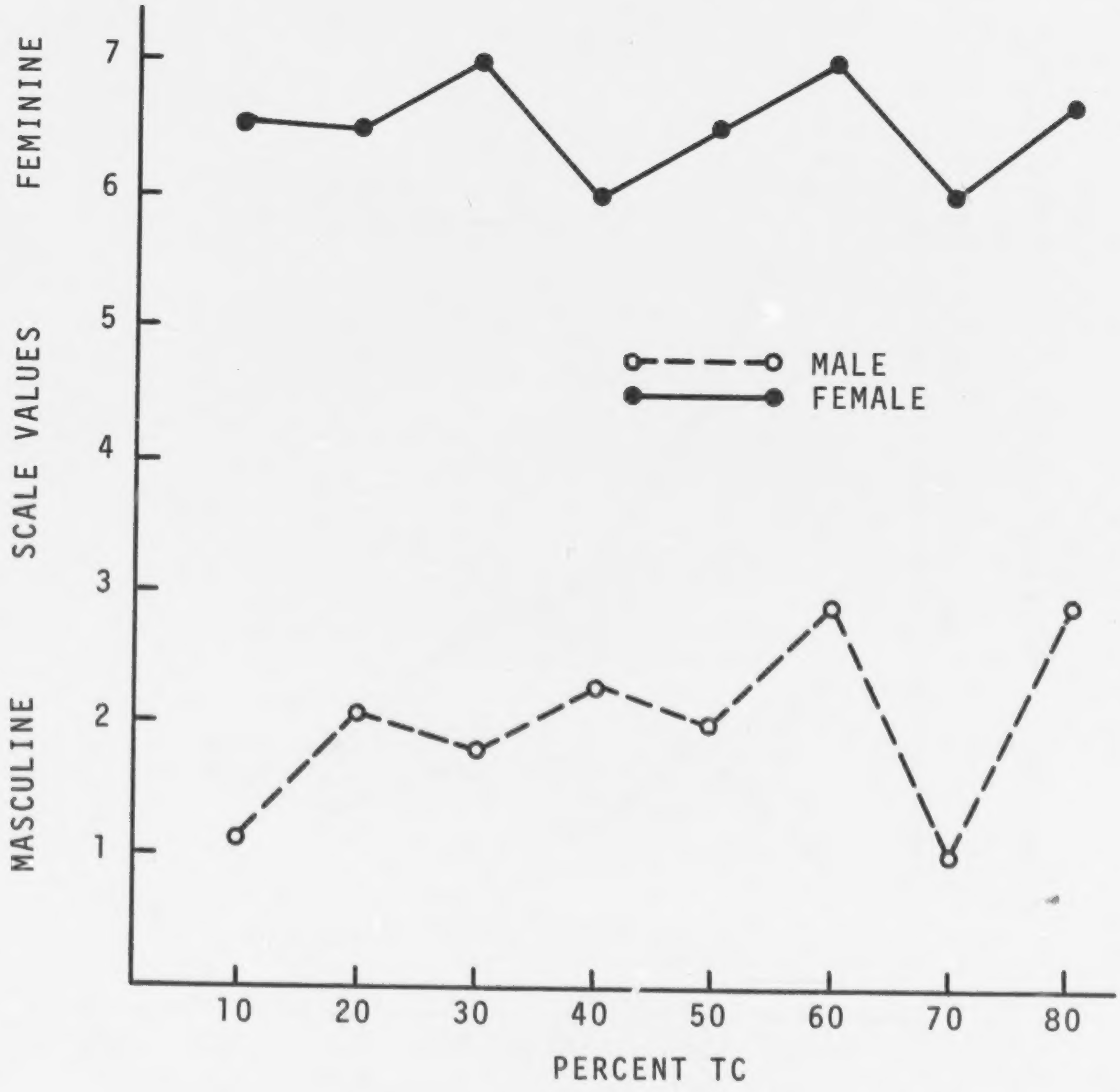


Figure 4.1. Graphic representation of listeners M scaled values of ratings of male and female TC vowels.

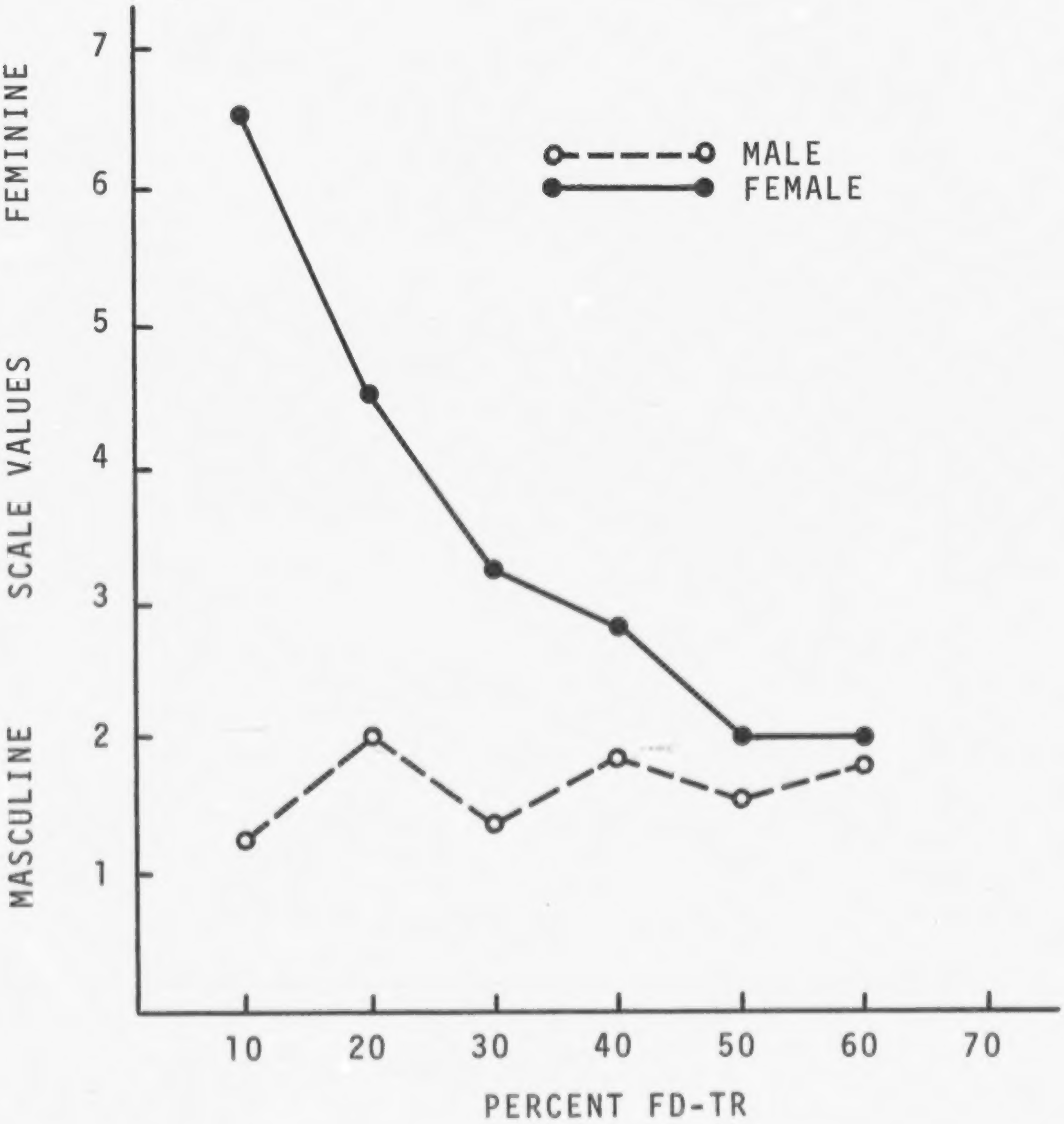


Figure 4. 2. Graphic representation of listeners M scaled values of ratings of male and female FD-TR vowels.

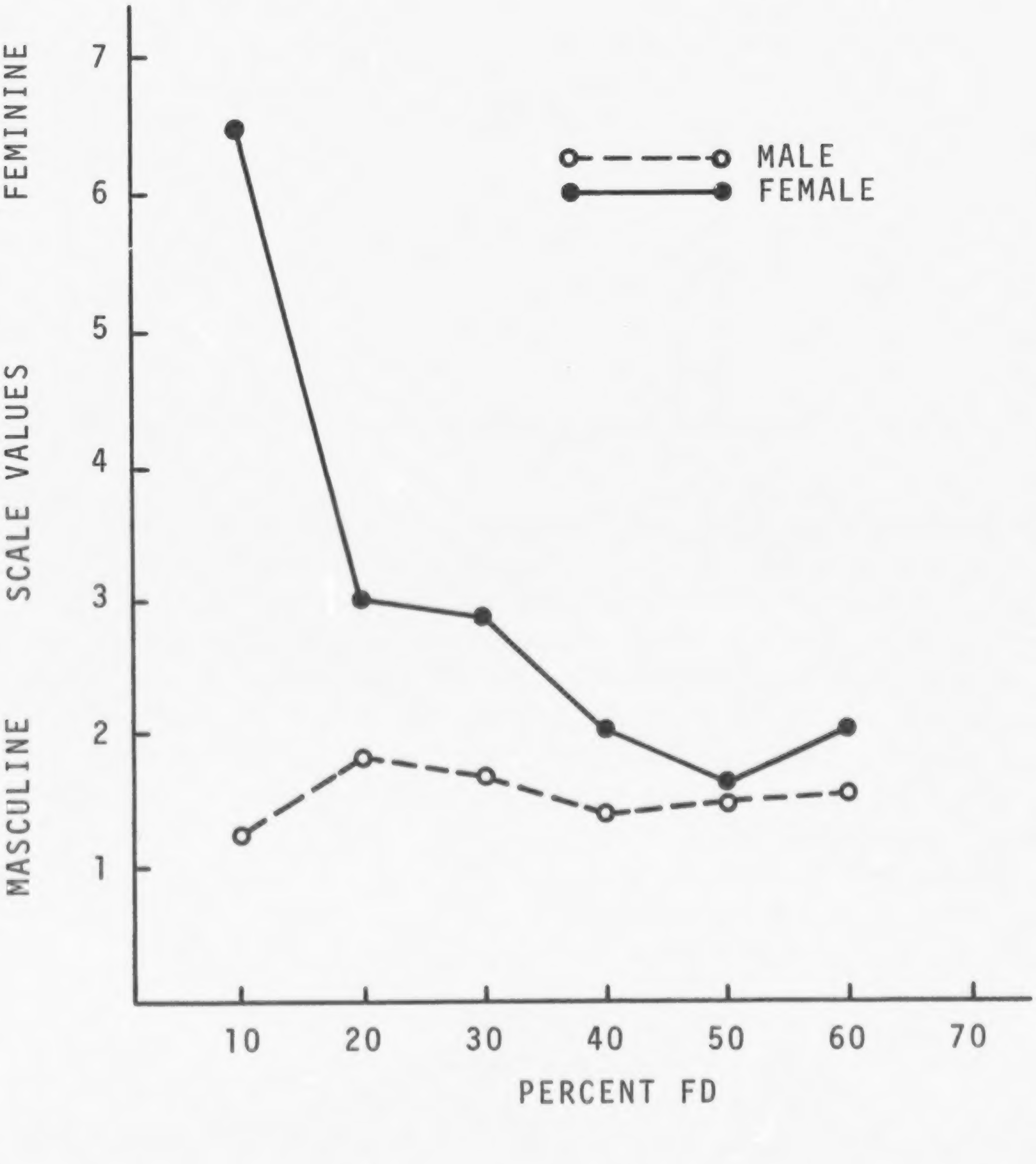


Figure 4. 3. Graphic representation of listeners M scaled values of ratings of male and female FD vowels.

TABLE 4.1

$\overline{M}$ VALUES OF SCALED MASCULINITY-FEMININITY OF LISTENER'S RESPONSES TO MALE AND FEMALE TIME-COMPRESSED (TC), FREQUENCY-DIVIDED (FD), AND FREQUENCY-DIVIDED TIME-RESTORED (FD-TR) VOWEL IN H-D CONTEXT FOR PHASE I

	TC		FD		FD-TR	
	Male	Female	Male	Female	Male	Female
0%	1.0	6.4	1.0	6.4	1.0	6.4
20%	1.7	6.4	1.7	4.2	1.5	3.0
30%	1.5	6.5	1.3	3.0	1.5	2.9
40%	1.9	6.0	1.4	2.6	1.1	1.5
50%	1.7	6.1	1.3	1.5	1.2	1.3
60%	2.2	6.2	1.3	1.5	1.3	1.7
70%	1.5	5.7	---	---	---	---
80%	2.0	5.8	---	---	---	---

The frequency-divided and frequency-divided time-restored conditions (Figures 4.2 and 4.3 respectively) show more profound experimental effects. From 20% on, under both conditions, the female appears to sound masculine. This initial effect is greater under the frequency-divided time-restored than frequency-divided condition. The frequency-divided time-restored curve is also steeper than the frequency-divided curve. Further, the frequency-divided time-restored maximum masculine rating for the female speaker is attained at 40%, whereas the frequency-divided maximum for the female is not attained until 50%. Finally, the frequency-divided condition maximum masculine rating for the female appears more stable than the frequency-divided time-restored maximum masculine rating for the female speaker.

Results--Phase II

The tentative results of this study suggest that a female speaker may not be preferred under conditions of frequency-divided and frequency-divided time-restored speech because of an effiminate perceptual quality after her speech has been processed.

Reliability of Results

Analyses of two scales were performed under Phase II: Masculine-Feminine, Intelligible-Unintelligible. Reliability coefficients computed for these data revealed an $r_{tt} = .98$ and $r_{tt} = .87$ respectively. Using the Silverman Estimation Method (Silverman, 1968), it was found that an additional five listeners would be required to raise the r_{tt} to .90 for the Intelligible-Unintelligible scale.

$\overline{M}$ Values of Ratings by Conditions

As expected, similar findings were obtained on the Masculine-Feminine scale in Phase II as were obtained in Phase I. One difference was that the maximum masculine rating for the female for frequency-divided and frequency-divided time-restored speech in Phase II was not reached until 60%. Table 4.2 and Figures 4.4 and 4.5 depict this information.

TABLE 4.2

$\overline{M}$ VALUES OF SCALED MASCULINITY-FEMININITY OF LISTENER'S RESPONSES TO MALE AND FEMALE FREQUENCY-DIVIDED (FD), AND FREQUENCY-DIVIDED TIME-RESTORED (FD-TR) VOWEL IN AN H-D CONTEXT FOR PHASE II

	FD		FD-TR	
	Male	Female	Male	Female
0%	1.40	6.5	1.4	6.5
20%	1.20	3.9	1.4	3.5
30%	1.20	3.1	1.2	3.3
40%	1.26	1.8	1.1	2.6
50%	1.26	1.8	1.1	1.7
60%	1.20	1.5	1.3	1.3

Regarding the Intelligible-Unintelligible scale values, the frequency-divided and frequency-divided time-restored conditions were both rated highly intelligible through the 20% condition. For both conditions the first major drop in intelligibility occurs at 30% for both sexes, the male speaker showing a steeper slope than the female. The data reveals the male speaker to be rated less intelligible than the female through the remaining compression levels for both conditions. The frequency-divided time-restored condition shows a more rapid decline in rated intelligibility than does the frequency-divided condition. For the frequency-divided time-restored condition the most dramatic drop occurs at 40% for the male, at 50% for the female. Although the frequency-divided condition reveals a more systematic decline in intelligibility, the frequency-divided time-restored condition appears to stabilize at higher compression

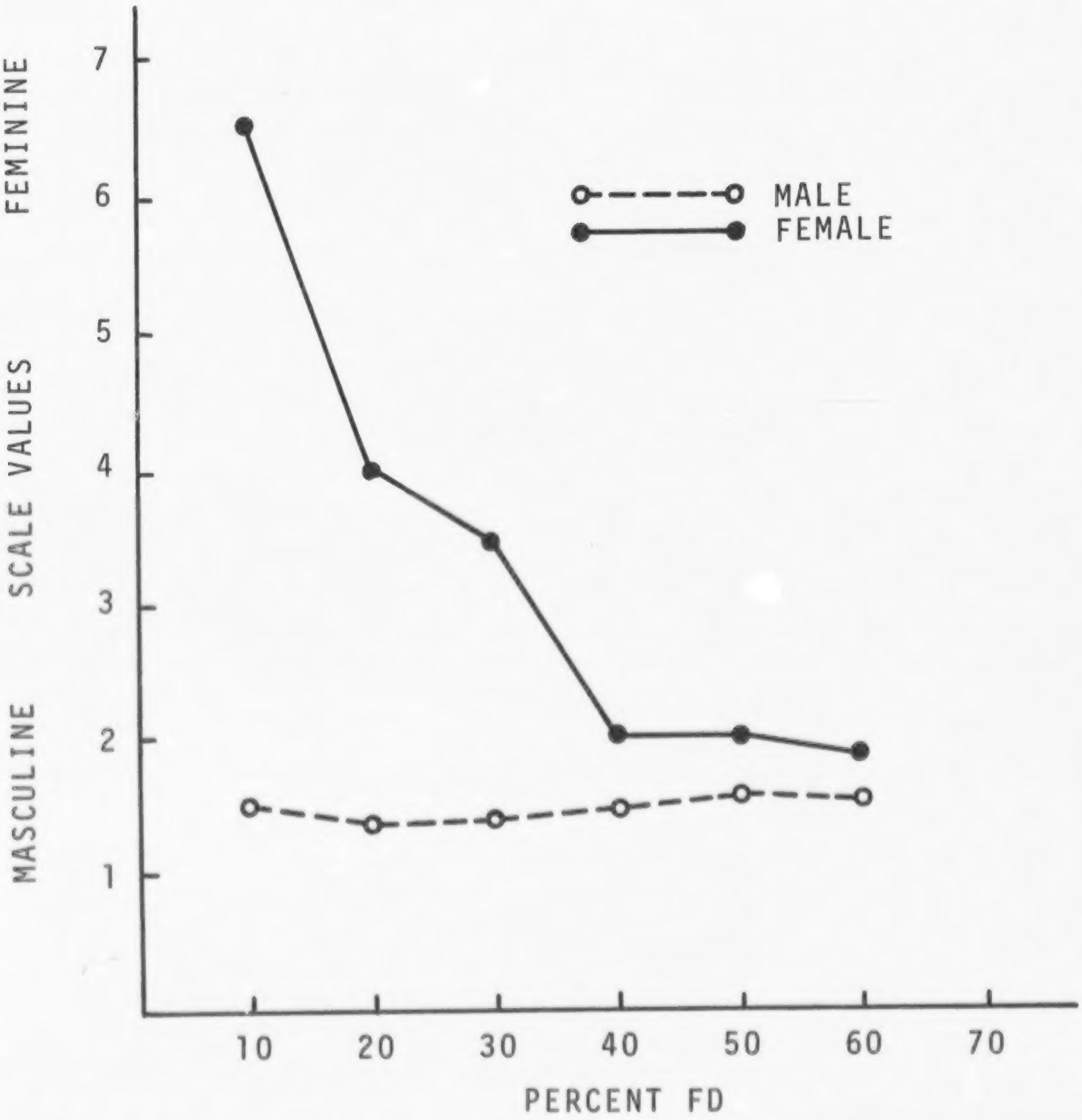


Figure 4.4. Graphic representation of listeners M scaled values of ratings of male and female FD vowels.

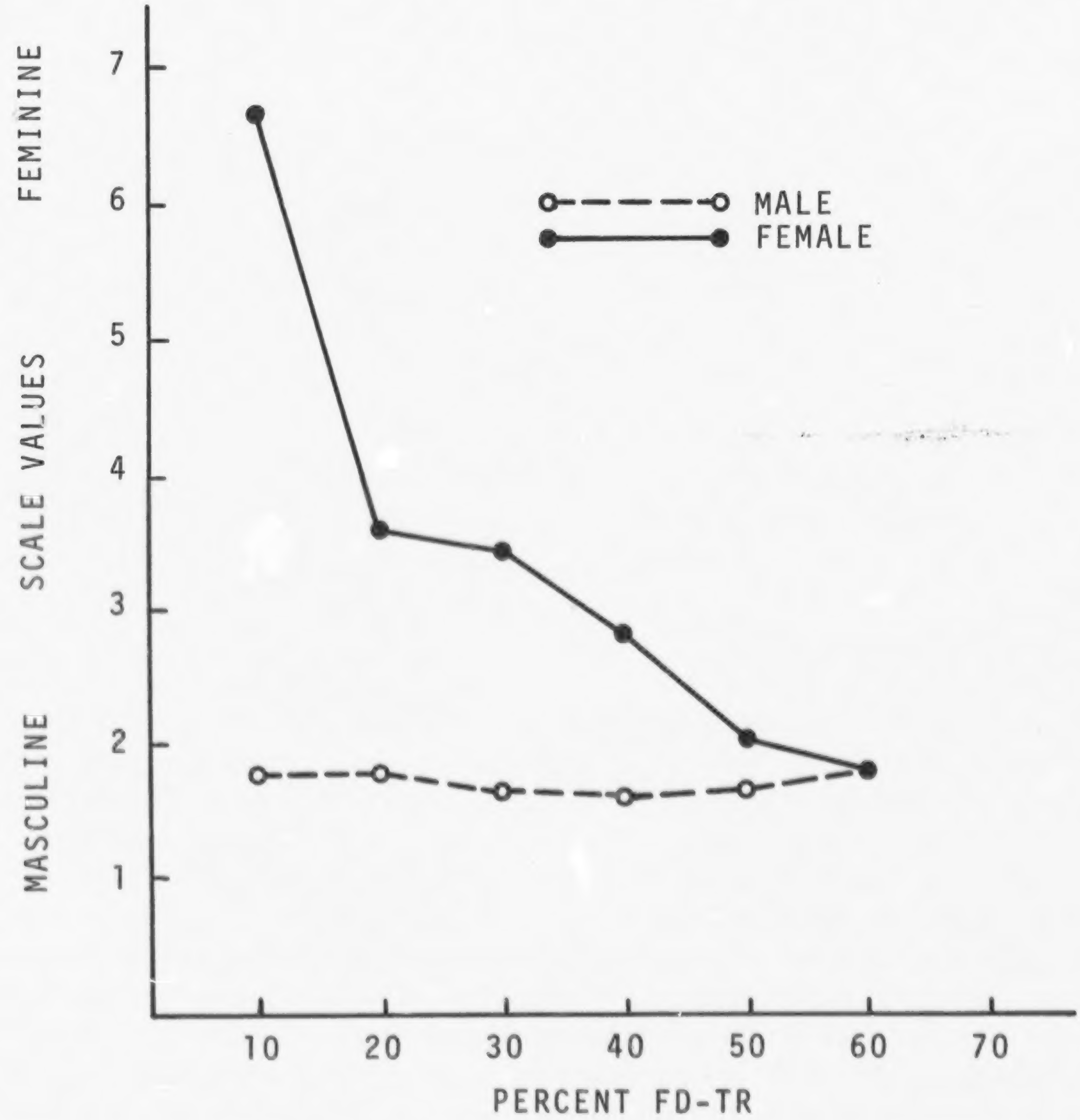


Figure 4.5. Graphic representation of listeners M scaled values of ratings of male and female FD-TR vowels.

condition (beyond 50% for both sexes). This data is summarized in Table 4.3 and Figures 4.6 and 4.7.

TABLE 4.3

$\overline{M}$ VALUES OF SCALED INTELLIGIBLE-UNINTELLIGIBLE OF LISTENER'S RESPONSES TO MALE AND FEMALE FREQUENCY-DIVIDED (FD) AND FREQUENCY-DIVIDED TIME-RESTORED (FD-TR) VOWELS IN AN H-D CONTEXT FOR PHASE II

	FD		FD-TR	
	Male	Female	Male	Female
0%	6.3	6.2	6.3	6.2
20%	6.7	6.2	6.1	6.4
30%	4.0	5.5	4.5	5.4
40%	3.1	4.4	2.7	4.8
50%	1.9	3.5	1.8	2.4
60%	1.3	2.5	2.3	2.4

Discussion

Time Compression

From the results it can be concluded that speaker sex identification under even extreme conditions of time-compressed speech tends to remain stable. The graphic depiction of the time-compressed ratings also tends to vary about the same for both sexes. Zemlin et al. (1968) concluded that intelligibility was not equivalent to preference, that is, what may be most intelligible may not necessarily be what is preferred. It was felt that a reason for this might be related to speaker sex identification under various conditions of time-compressed speech. The question is still to be resolved as to the essential differences between the Foulke (1966c) findings and those of Daniloff et al. (1968a). Further analysis of several of the other semantic differential scales used in this study is currently underway.

Frequency-Divided and Frequency-Divided Time-Restored

The results agree with Daniloff et al. (1968a) and Klumpp and Webster (1961) in that the female frequency-divided and frequency-divided time-restored speech is more intelligible than the male frequency-divided and frequency-divided time-restored speech. Further agreement with the

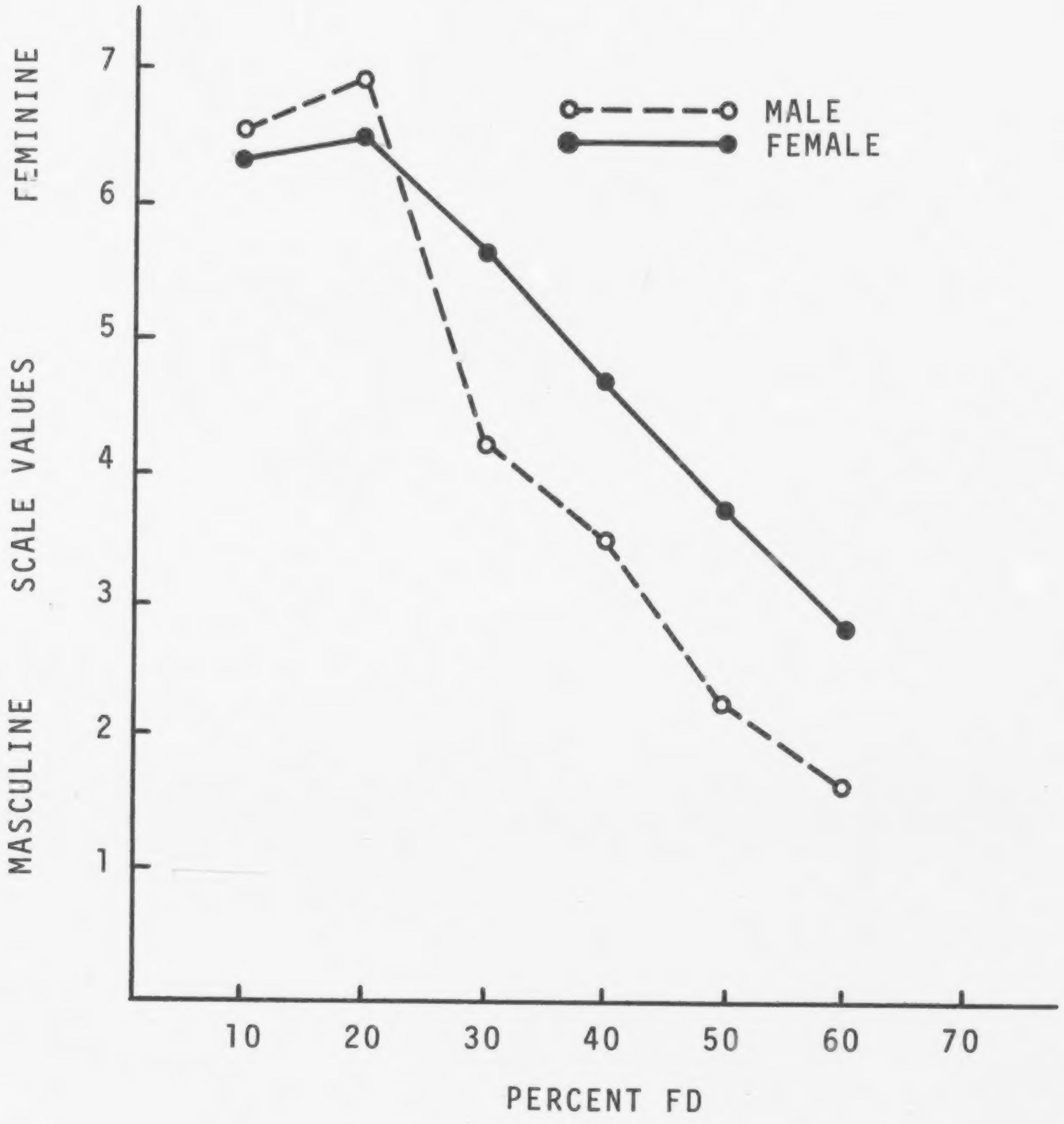


Figure 4. 6. Graphic representation of listeners M scaled values of ratings of male and female FD vowels.

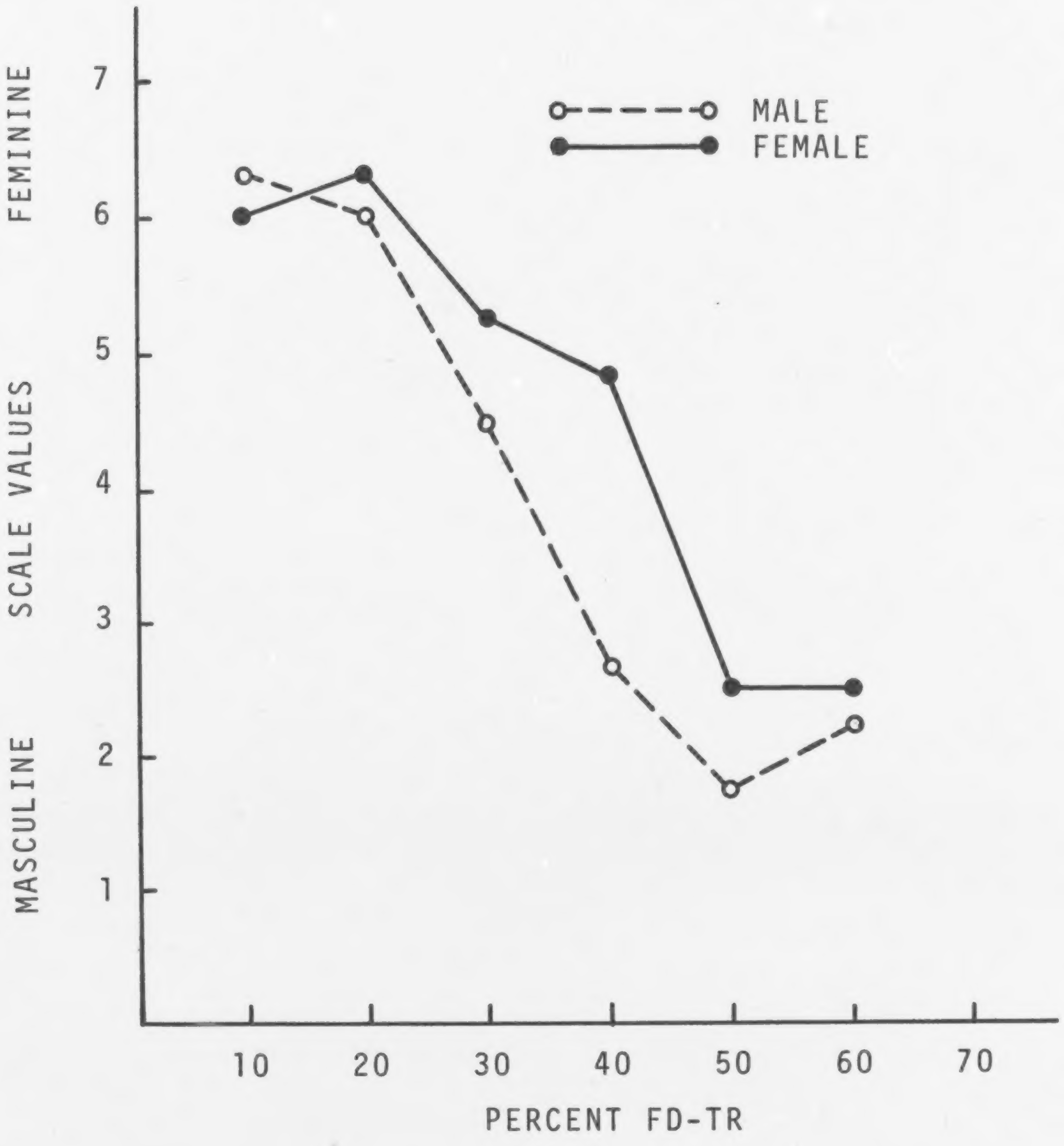


Figure 4. 7. Graphic representation of listeners M scaled values of ratings of male and female FD-TR vowels.

Daniloff et al. (1968a) study is seen in that both studies reveal the first major decline in intelligibility to be about 30% distortion. Also, the frequency-divided time-restored speech revealed a rapid initial decline in both studies, especially for the male speaker. Finally, both studies reveal the most dramatic drop for the male frequency-divided time-restored speech to be at 40%, for the female at 50%.

The agreement between these studies relative to the Intelligible-Unintelligible scales suggests that a listener is able to judge adequately what is intelligible to him, and that this judgment would be highly correlated to what would be revealed by traditional intelligibility tests.

The conflicting results between the Daniloff et al. (1968a) study and this study, that the female frequency-divided and frequency-divided time-restored speech is more intelligible, and the Tiffany and Bennett (1961) study, which showed a male preference, can be explained by the results of this study. Apparently the female distorted speech begins to take on a psychological male-like component, whereas the male speaker, as expected, tends to remain stable. There is no social decision to be made with respect to his distorted speech. These findings would support the contention that phonetic quality and phonemic quality are not equivalent. Further, phonetic quality may be based on a fixed vowel hypothesis, whereas phonemic quality may be related to a relative vowel hypothesis.

Further analyses of this data are being carried out. Further, physical measurements, such as those performed by Terango (1966), are being performed on all four speakers in order to physically account for the gradual shift of the female to male frequency-divided and frequency-divided time-restored speech. It is suspected that the female frequency-divided and frequency-divided time-restored speech will reveal that the M rate of pitch change during inflection will decrease with increased distortion, as revealed by Terango (1966) when he studied rated effeminate voices.

Finally, the results of the Like-Dislike scale should shed substantial light upon the preference/intelligibility controversy.

There appears little doubt that if time-compressed, frequency-divided, and frequency-divided time-restored processed speech is to be used educationally, consideration must be given to more than simply intelligibility. What an individual likes (prefers) to listen to may have significant bearing on his progress.

ACKNOWLEDGMENTS

This investigation was supported in part by a Public Health Service Training Grant TI-NG-54-79 from the National Institute for Neurological Diseases and Stroke, and in part by U.S. Public Health Service Training Grant USPH-1-T1-DE-169 from the National Institute for Dental Research. The Authors extend gratitude to Allan Bird for his assistance in the analysis of this study.

APPENDIX A

INSTRUCTIONS FOR SCALING STUDY ON THE PERCEPTION OF DISTORTED SPEECH

The purpose of this study is to study the feelings of people to various types of speech. We hope to do this by having the people judge the speech they hear against a series of descriptive scales. In taking this test, please make the judgment on the basis of how you feel about the speech signals you are to judge. On the dittoed sheet you will find seven different scales. I will play a recorded tape. You will hear vowels in an h-d context. There are 41 sets of 11 vowels per set. Between each set of 11 vowels there is a silence of about 1 minute. During this silence following each set, you are to rate the set on the seven scales, in order.

Here is how you are to use the scale:

If you feel that the set of words you heard is very closely related to one end of the scale, you should place your checkmark as follows:

fair __X__: _____: _____: _____: _____: _____: _____: unfair

OR

fair _____: _____: _____: _____: _____: _____: __X__: unfair

If you feel that the set of words in quite closely related to one or the other end of the scale (but not extremely), you should place your checkmark as follows:

fair _____: __X__: _____: _____: _____: _____: _____: unfair

OR

fair _____: _____: _____: _____: _____: __X__: _____: unfair

If the set of words seems only slightly related to one side or the other side (but not really neutral), then you should check as follows:

fair _____: _____: __X__: _____: _____: _____: _____: unfair

OR

fair _____: _____: _____: _____: __X__: _____: _____: unfair

The direction toward which you check, of course, depends on which of the two ends of the scale seem most characteristic of the set of words you are judging.

If you consider the set to be neutral on the scale, both sides of the scale equally associated with the set you are judging, then you should place your checkmark in the middle space:

fair _____: _____: _____: __X__: _____: _____: _____: unfair

IMPORTANT:

(1) Place your checkmarks in the middle of the spaces, not on the boundaries:

_____: _____: __X__: _____: _____: _____: _____X:
(this) (not this)

(2) Be sure to check every scale for every set of words-- do not omit any.

(3) Never put more than one checkmark on a single scale.

(4) Remember, there's only about a minute between sets, so work accurately but rapidly.

Sometimes you may feel as though you've had the same set of words before on the test. This will not be the case, so do not look back at previous rating sheets. Do not try to remember how you checked previous items on each scale: make each item a separate and independent judgment. Do not worry or puzzle over individual items. It is your first impression, the immediate feeling about the sets of words we want. On the other hand, do not be careless because we want your true impressions.

Are there any questions?

Do not begin your ratings until the set ends.
Three beeps mean get ready, one beep means the end of the set.
This is not a test of intelligibility.

CHAPTER V

DICHOTIC SPEECH-TIME COMPRESSION

Sanford E. Gerber and Robert J. Scott*

In general, the Fairbanks procedure (Figure 5.1) or its German equivalent has been the method of choice for various applications. The main difficulty with time compressing speech in this way is that it depends for its compression upon the discarding of information. If the intelligibility is less than that achieved uncompressed, it is probably due to the loss of information. Scott (1965, 1967b), making this observation, hypothesized that restoring the information should restore the intelligibility. We have now completed a series of studies to verify this hypothesis.

Dichotic Compression

Scott's (1965) procedure is called "dichotic" speech-time compression. Recall that in the Fairbanks procedure the signal (and hence, the information) in the discard interval is not recoverable, so could not be made available to the listener. The differences between "diotic" speech-time compression (i.e., Fairbanks' method) and "dichotic" speech-time compression (i.e., Scott's method) are shown in Figure 5.2. The diotically produced tape has one track which contains only the (imaginary) odd-numbered segments, and these continuous segments are heard in both ears. The dichotically produced tape has two tracks: one track is identical to the diotic tape and is played to one ear only; the second track contains only the (imaginary) even-numbered segments, and this track is played only to the other ear. Notice that the first track is delayed a bit so that it is offset in time with respect to the second track. The second time segment no longer follows the first, but overlaps it in time. The significance of the amount of overlap remains to be investigated, but for all these experiments it has been 50% with respect to either segment.

---

*Dr. Sanford E. Gerber is Assistant Professor of Speech and Disorder of the Audiology Laboratory at the Speech and Hearing Center of the University of California, Santa Barbara, California. Dr. Robert J. Scott is a consultant with the U.S. Government, Washington, D.C.

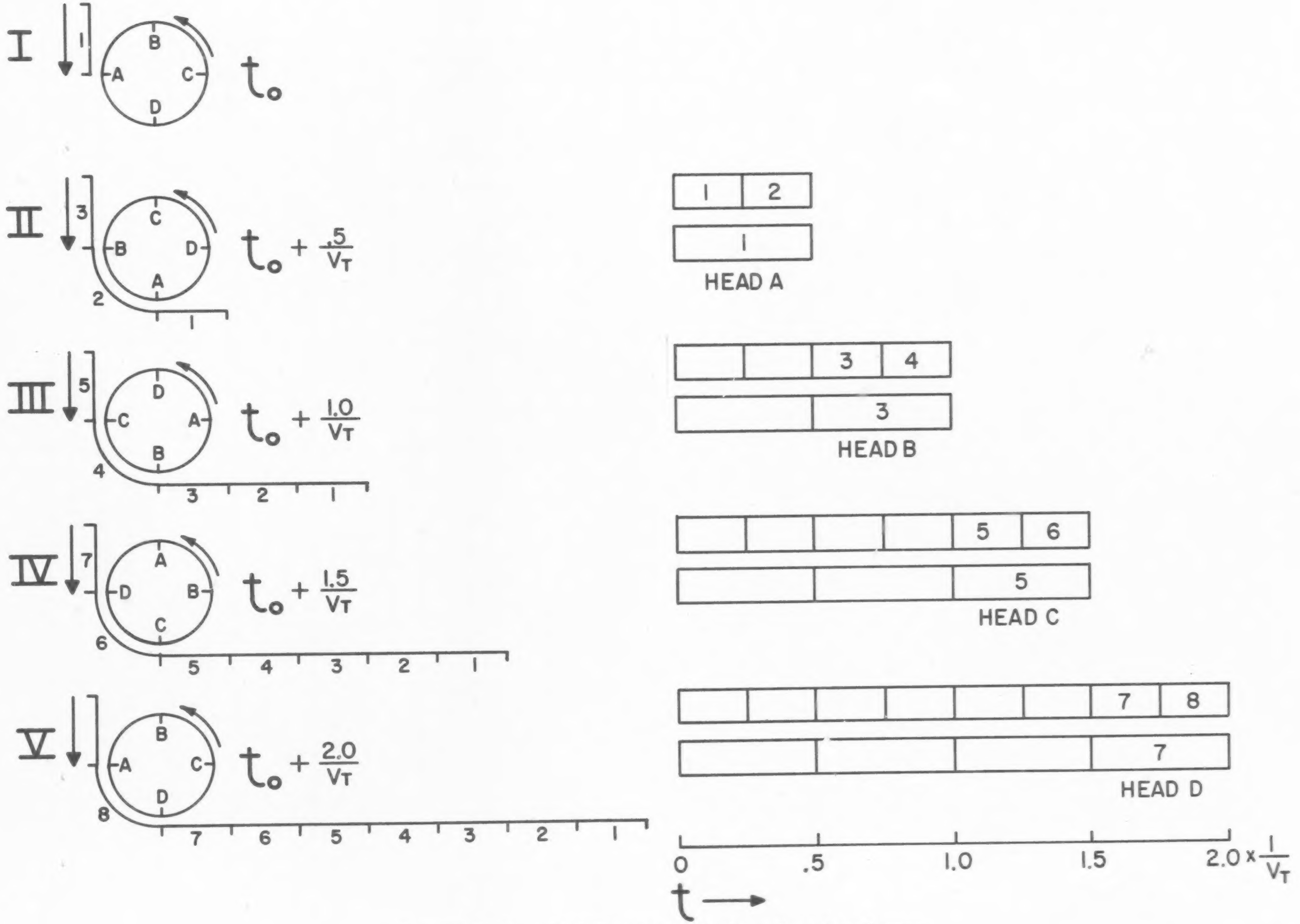

I
II
III
IV
V
A
B
C
D
t_o
$t_o + \frac{.5}{V_T}$
$t_o + \frac{1.0}{V_T}$
$t_o + \frac{1.5}{V_T}$
$t_o + \frac{2.0}{V_T}$
HEAD A
HEAD B
HEAD C
HEAD D
0
.5
1.0
1.5
$2.0 \times \frac{1}{V_T}$
t →

Figure 5.1. Fairbanks' rotating head scheme.

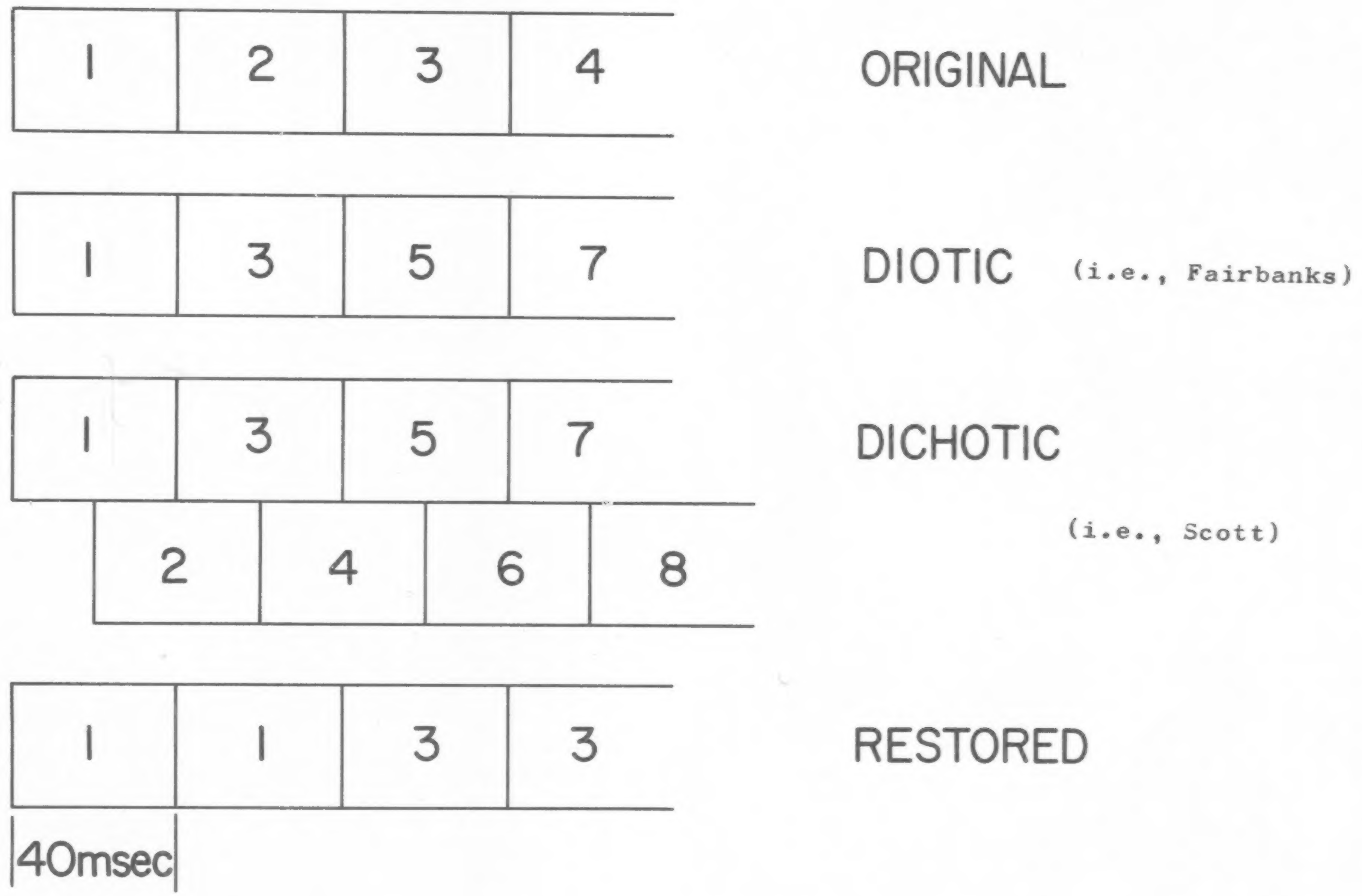

1
2
3
4
ORIGINAL
1
3
5
7
DIOTIC (i.e., Fairbanks)
1
3
5
7
2
4
6
8
DICHOTIC
(i.e., Scott)
1
1
3
3
RESTORED
40msec

Figure 5.2. 2:1 compression.

To create dichotic speech-time compression a hybrid computer system has been used (Gerber, in press; Hogan & Scott, 1963). For the research reported here, two different (but very similar) hybrid computers have been employed. Figure 5.3 shows the hybrid system used for the experiments up to the last one, and Figure 5.4 shows the system used in the latest experiment. The input/output apparatus is essentially the same in both cases. The older system uses a PDP-1 digital computer, while the newer one employs a PDP-7. The PDP-1 is a somewhat larger but slower machine than the PDP-7; both are manufactured by the Digital Equipment Corporation of Maynard, Massachusetts. The analog portion of the hybrid is a Pace TR-10 analog computer associated with the PDP-1 or an EAI 8800 in the case of the PDP-7. The analog computers were made by Electronic Associates, Inc. of Long Branch, New Jersey. The writers have been very fortunate to have had these hybrid computer systems available for this research.

Experiment I

Our first investigation of the intelligibility of dichotic speech-time compression dealt with the differences between dichotic and diotic for each of three compression ratios. The results of that study have been reported (Gerber, 1968) and need only be summarized here.

The stimuli used in all the intelligibility experiments were Fairbanks' recordings of the rhyme test words (cf. Fairbanks, 1958). The recordings of the rhyme tests were input from the tape playback via the analog computer interface to the analog-to-digital converter which put the digitized speech onto magnetic tape. Then, under operator control, the computer time compressed the digitized speech and wrote this version onto another magnetic tape. When the compression process had been completed, the compressed digital tape was output via the digital-to-analog converter onto audio tape. In this way all 250 items of the Fairbanks recordings were compressed and dichotomized.

In the first experiment, we used three different compression ratios (R_c = 2:1, 3:2, and 4:3) and three different discard intervals (I_d = 30, 40, and 50 milliseconds [msec.]). Twenty listeners were employed. For all listeners, the dichotic signals were more intelligible than their diotic counterparts. Combining across both compression ratio and discard interval, it was found that dichotic listening led to higher intelligibility scores than diotic listening. For this aggregate of dichotic signals, the average intelligibility exceeded 97%; while, for the aggregate of diotic signals, the average intelligibility was just over 93%. This difference was significant beyond the 0.01 level. This means that the dichotic version did, indeed, restore the intelligibility; and presumably via the restoration of the otherwise discarded information.

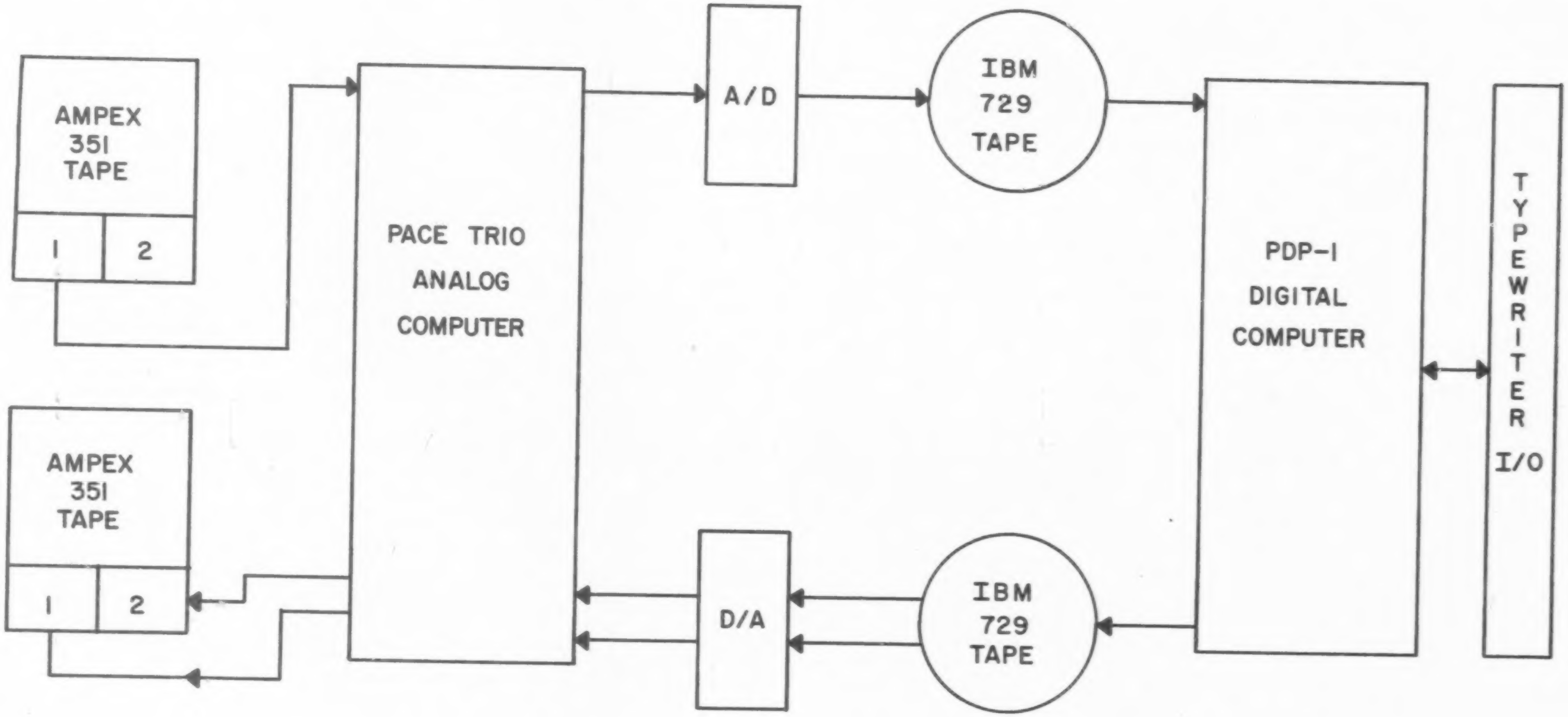

AMPEX
351
TAPE
1
2
AMPEX
351
TAPE
1
2
PACE TRIO
ANALOG
COMPUTER
A/D
D/A
IBM
729
TAPE
IBM
729
TAPE
PDP-1
DIGITAL
COMPUTER
TYPEWRITER
I/O

Figure 5.3. Hybrid computer system number one.

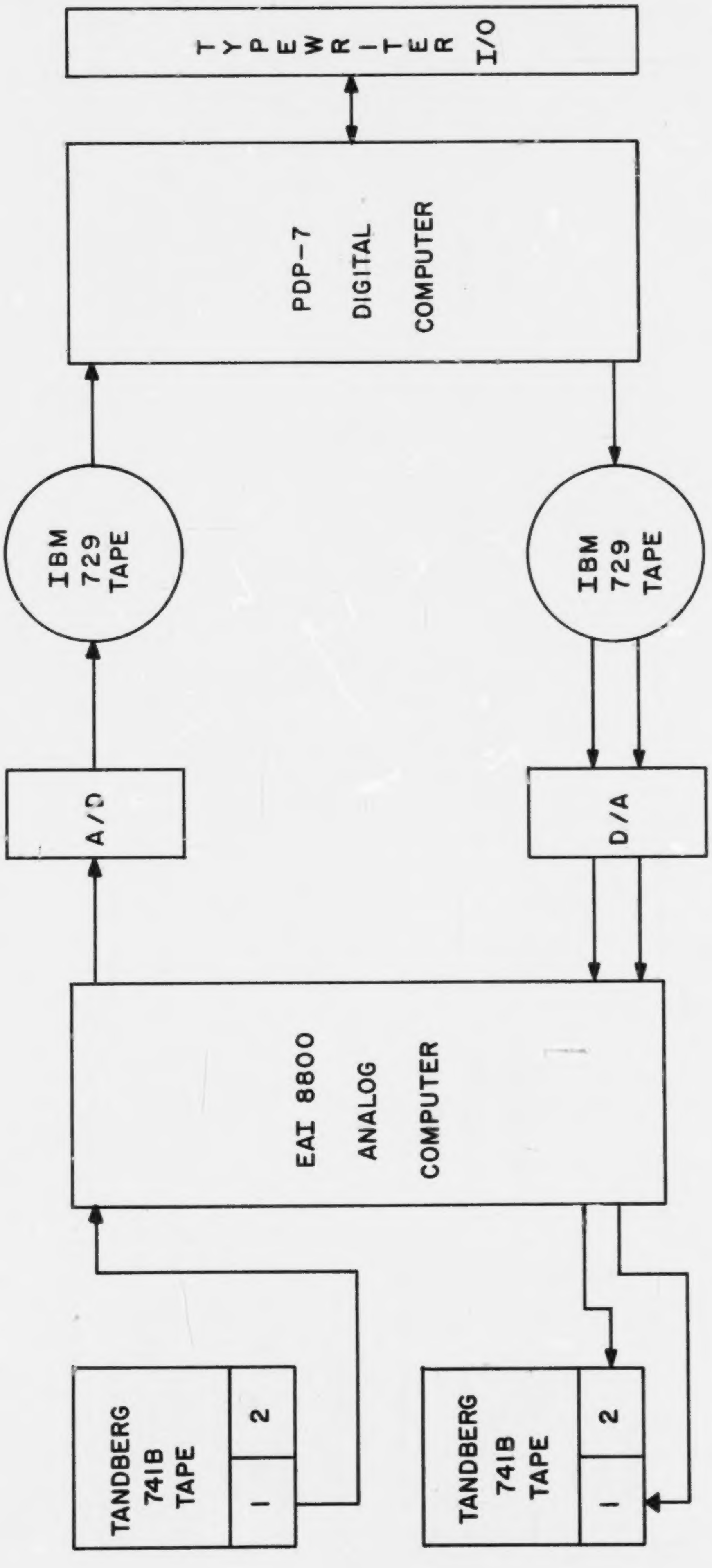

TYPEWRITER I/O
PDP-7 DIGITAL COMPUTER
IBM 729 TAPE
IBM 729 TAPE
A/D
D/A
EAI 8800 ANALOG COMPUTER
TANDBERG 741B TAPE
1
2
TANDBERG 741B TAPE
1
2

Figure 5.4. Hybrid computer system number two.

There were some other interesting findings from this first experiment. We found no significant differences among the three compression ratios when discard interval was not considered. Therefore, we could conclude that the dichotic restoration of information was good to at least double normal speed. Moreover, virtually no intelligibility was to be gained by minimizing the amount of compression, for example, to only 4:3 or 3:2.

We did find differences with respect to discard interval. The discard interval of 50 msec. was significantly (at the 0.10 level) less intelligible than the others, which did not differ significantly from each other. It seems, then, that a discard of 50 msec. is in some way "too long." When the information was restored, that is when dichotic was compared with diotic, it was seen that the restoration was significant for the 50 msec. discard interval. The difference between dichotic and diotic for this discard interval was nearly 9%. Therefore, given a diotic signal, 50 msec. of information is too much to miss at one time.

Although many early experiments support the use of a discard interval of about 40 msec. for all compression ratios, we feel that for isolated words intelligibility can be significantly increased by using a discard interval as short as 15 to 20 msec. depending on the amount of compression and the average fundamental frequency of the speaker. For continuous speech, shorter discard intervals have been avoided primarily because of the annoying effect of the interruption frequency. As those of us working with computer-compressed speech have experienced, the use of sampling intervals which do not preserve at least one complete and continuous voicing period injects an artificial monotone pitch, the frequency of which is inversely proportional to the sampling period.

In general, we concluded from this first experiment that time-compressed speech (up to double speed) is highly intelligible anyway, and restoration of the information by Scott's dichotic technique is not only feasible but desirable. His scheme of dichotic speech-time compression restores the intelligibility of time-compressed speech essentially to its uncompressed level.

Experiment II

The results of Experiment I were very encouraging, but left some questions unanswered. Speech compressed by means of discarding segments leads to more listening possibilities than were investigated in Experiment I. Reference to Figure 5.5 will show that there are four listening possibilities. What we had called "diotic" referred to the fact that the signal was heard with both ears when there was only one signal. To describe this, in the second experiment we decided to label this condition "Unitary-

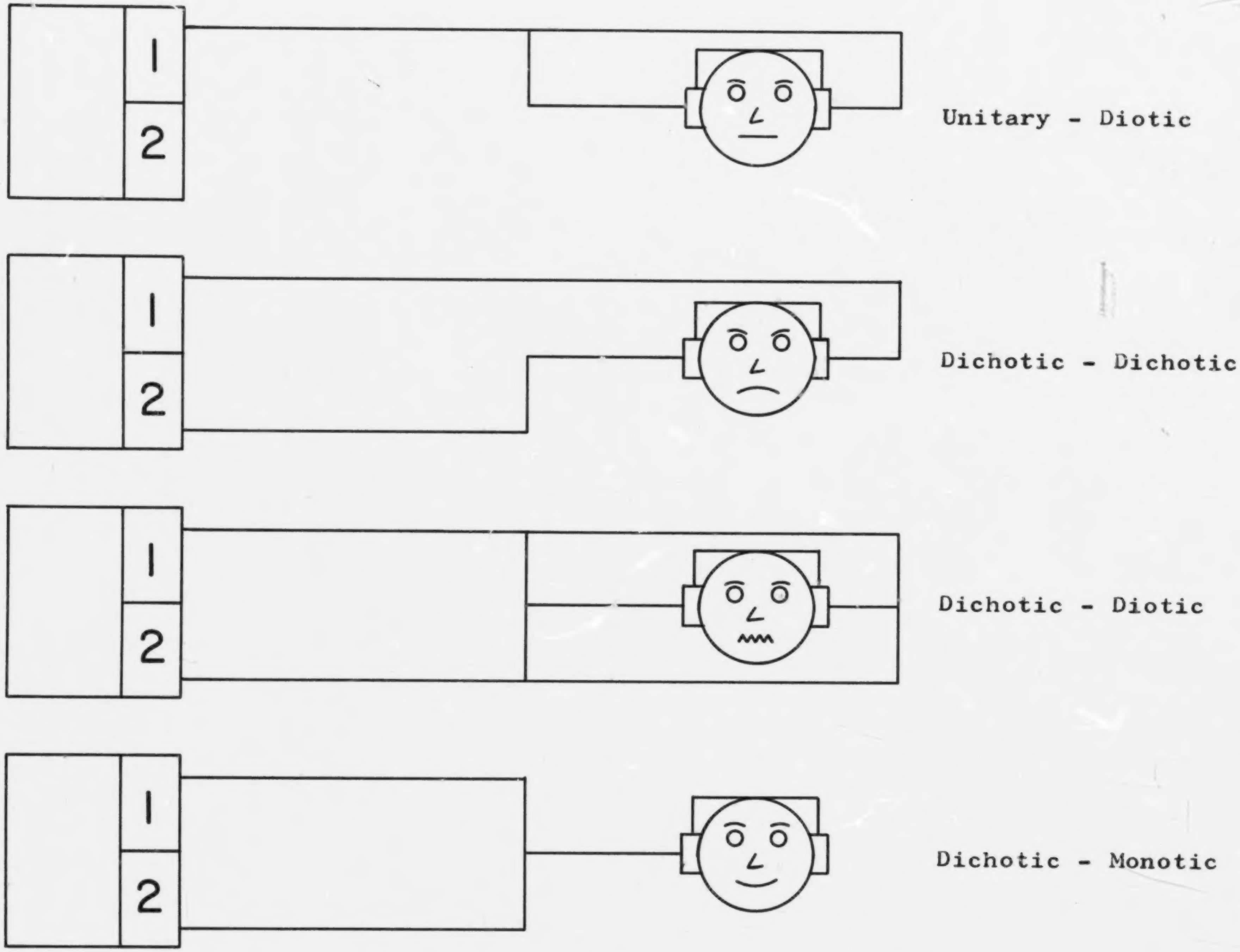

1
2
Unitary - Diotic
1
2
Dichotic - Dichotic
1
2
Dichotic - Diotic
1
2
Dichotic - Monotic

Figure 5.5. Listening possibilities.

Diotic," meaning one signal in two ears. The other possibilities using this scheme are: Dichotic-Diotic, Dichotic-Monotic, and Dichotic-Dichotic. In these terms, we had looked in the first experiment only at Unitary-Diotic (one signal in both ears) and Dichotic-Dichotic (two signals, one in each ear). We were, therefore, unable to tell whence came the apparent improvement: from the dichotomizing of the signal, or from the dichotomizing of the listener.

The purpose of Experiment II was to determine the necessity of listening dichotically to dichotic signals. Perhaps one ear could process dichotic time-compressed speech as well as two since all the information would have been restored in this case, too. The results of this investigation are also in the literature (Gerber, 1969). It was found that dichotic signals heard diotically were not more intelligible than when heard monotically; there was no significant difference between monotic and diotic listening conditions when the signal was dichotic. Moreover, no preference for ear was observed in the monotic condition. Using a dichotic signal it seems sure (and not at all surprising) that intelligibility would be superior if the listening condition were also dichotomized. That is to say, the highest intelligibility results when the dichotic signal is presented one track to each ear. If both tracks are presented to one or to both ears, intelligibility suffers. Furthermore, if only one track is used (in one or both ears) intelligibility suffers.

Experiment II, like Experiment I, revealed a significant difference between 40 and 50 msec. discard intervals. Again, intelligibility with discards of 40 msec. was greater than with discards of 50 msec. with the compression ratio fixed at 2:1. The data of Experiment I caused us to decide that ratios less than 2:1 were no longer interesting.

These two experiments led us to raise a question which has been asked many times over the years that this process has been investigated. Why is time-compressed speech less intelligible than uncompressed? Is the loss of intelligibility due solely to the loss of information? Or is it due to excessive rate demands upon the listener? Or both?

Experiment III

It seemed reasonably clear after the first two experiments that the loss of intelligibility must be due to the loss of information and not due to the speed being too demanding for processing. The dichotic signal, wherein all the information is present, was always more intelligible than the diotic signal at the same rate. However, what happens when the speed is more than doubled? If the compression ratio is greater than 2:1, it is not possible to recover all the discarded information (since we have only two ears), but it is possible to recover some of it.

If the intelligibility of time-compressed speech at greater than double the normal speed is enhanced by dichotomizing, then the loss of intelligibility must be attributable to the loss of information. Furthermore, it is possible to restore the time without restoring the information. If this time-restored version is not more intelligible than the compressed version, then one must conclude that the time demands are not excessive. Anyway, since the dichotic speech-time compression at rates up to double the original was shown to suffer no important loss of intelligibility, one wants to know how much compression will cause intelligibility to diminish significantly.

This third experiment, then, was intended to answer these questions. For this experiment we prepared tapes of the Fairbanks Rhyme Test (in the same manner as previously but with the newer hybrid computer) at a compression ratio of 3:1. Figure 5.6 shows the imaginary segment numbers available when these high compressions are recorded dichotically. It is seen there that all of the information is not restored by dichotomizing. By definition, a ratio of 3:1 diotic contains one-third of the information in the original signal; dichotomizing restores another third. Dichotic compression at 3:1, then, contains two-thirds of the information of the uncompressed signal.

Experiment III presented three different modes of compression to the listeners. All the modes were at 3:1 with a discard interval of 40 msec. The three modes were: dichotic, diotic, and time-restored. Reference again to Figure 5.6 reveals that in order to restore the time but not the information it is necessary to repeat the same segments used in the diotic mode. If the compression ratio is 3:1, each segment is repeated three times and only one-third of the segments are used.

The decision to restore the original time frame by repeating the diotic (single file) compressed speech as in Figure 5.6 perhaps was not a wise one. The results may have been more intelligible for the time-restored words had the restoration been done dichotically. Preliminary data from a current experiment suggest that 3:1 dichotically restored words presented dichotically will prove to be more intelligible than 3:1 dichotically compressed words. We initially believed that diotically time-restored isolated words would be more intelligible than diotically time-compressed words because of earlier experiments in restoring the time of continuous speech. We now feel that repeating already distorted sampling intervals in order to time-restore isolated words tends to increase listener confusion, whereas this distortion tends to be less effective in continuous speech.

Figure 5.7 shows the results of this investigation compared with those of Experiment I. Most of our hypotheses have been verified by Experiment III. Again, dichotic processing made a significant (< 0.01)

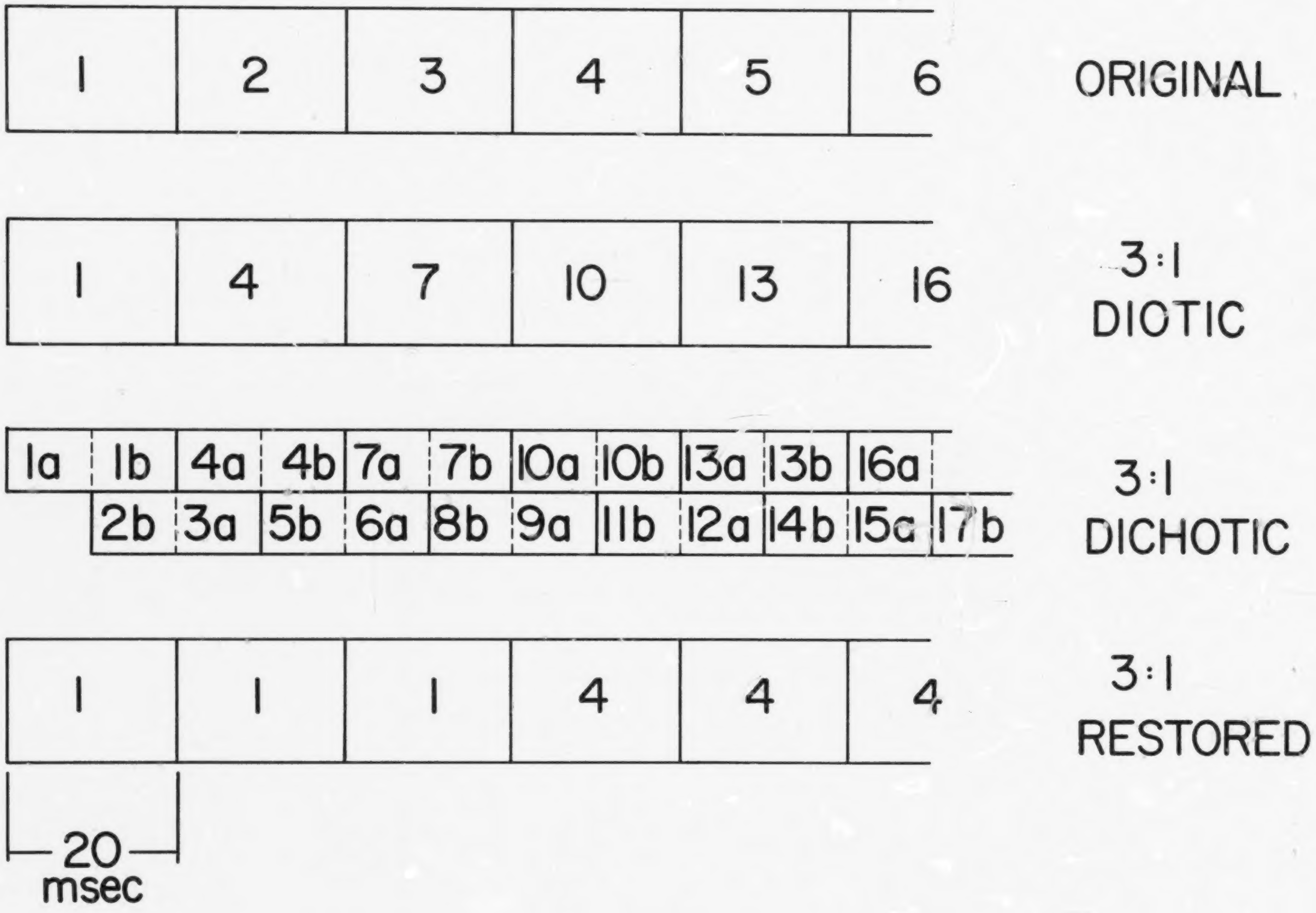

1
2
3
4
5
6
ORIGINAL
1
4
7
10
13
16
3:1
DIOTIC
1a
1b
4a
4b
7a
7b
10a
10b
13a
13b
16a
2b
3a
5b
6a
8b
9a
11b
12a
14b
15a
17b
3:1
DICHOTIC
1
1
1
4
4
4
3:1
RESTORED
20
msec

Figure 5.6. 3:1 compression.

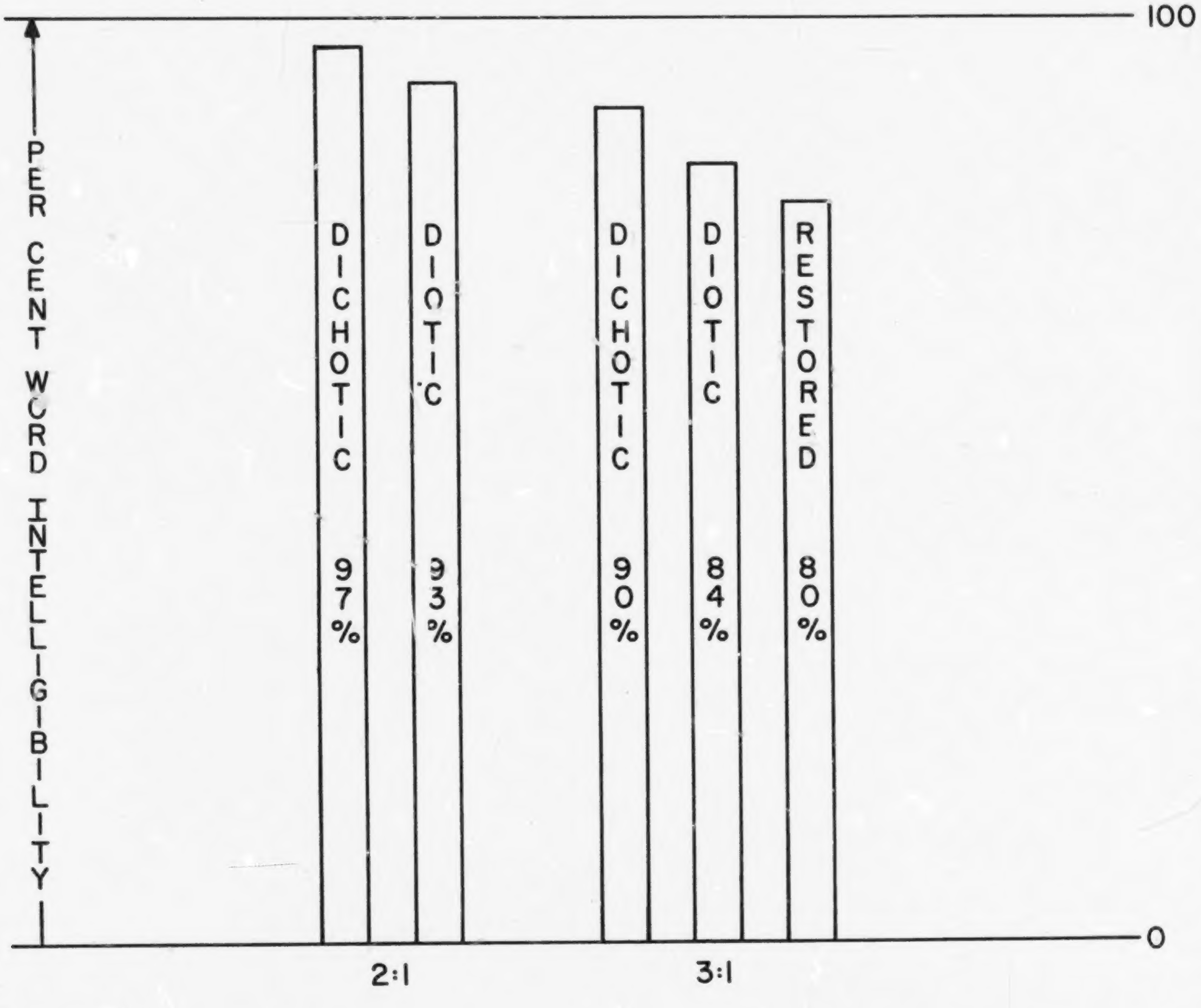


Figure 5.7. Results

improvement over the intelligibility possible diotically; this improvement was over 5%. Of even greater interest was the lack of improvement resulting from time restoration. To restore the normal time by repeating segments results in a very peculiar sounding signal; so peculiar, in fact, that its intelligibility is significantly poorer than even diotic (< 0.05) and much, much poorer than dichotic (< 0.001). So, time restoration is not the answer; at least, not time restored in this way because it introduces another kind of distortion.

We have not really resolved the issue of whether the loss of intelligibility is due to the lack of information or to the press of time. It is true that the dichotic signal even at 3:1 is really quite intelligible, which continues to suggest that the problem rests in the information and not the speed. The fact that the time-restored signal was so poor lends some credence to this hypothesis, but the restored speech has peculiarities of its own. We have resolved, however, that listeners can process speech at a very high rate even when there is a lack of information. The next study to be done may be the one which resolves this issue. A 4:1 dichotic signal contains as much information as a 2:1 diotic signal. If the losses are due solely to losses of information, then these should have equal intelligibility. If not, then 4:1 may be "too fast." Meanwhile, we find that we are well within human auditory processing time capabilities even at triple normal speed. The premise that loss of intelligibility in time-compressed speech is due primarily to the inability of the listener to process the speech at the higher rate is certainly an attractive one. For if it were true, it would suggest the possibility of training subjects in high-speed listening.

CHAPTER VI

A COMPARISON OF "DICHOTIC" SPEECH AND SPEECH COMPRESSED BY THE ELECTROMECHANICAL SAMPLING METHOD*

Emerson Foulke and E. McLean Wirth**

Recorded speech may be compressed in time by reproducing a succession of periodic, time-abutted samples of the original recording. If the durations of the samples eliminated from such a reproduction are brief enough so that no critical feature of a speech signal can, by accident of sampling, fall entirely within a discarded sample, the result is time-compressed, intelligible speech that is not altered with respect to vocal pitch or quality.

Such sampling may be accomplished manually (Garvey, 1953b), by cutting a recorded tape into segments, discarding some of the segments, and splicing the remaining segments together again. It may be accomplished more conveniently by a tape reproducer of the type described by Fairbanks, Everitt, and Jaeger (1954). Devices of the Fairbanks type reproduce periodic, time-abutted samples of a recorded tape and, as before, the result is time-compressed, intelligible speech, without distortion in vocal pitch or quality (Foulke, 1969).

---

*The research described in this report was also reported by the junior author in her senior thesis, submitted to the Webster College, St. Louis, Missouri, 1968. This report also appears as Chapter III in The Comprehension of Rapid Speech by the Blind: Part III, Final Progress Report, March 1, 1964 - June 30, 1968, Project No. 2430, Grant No. OE-4-10-127, U. S. Department of Health, Education, and Welfare, Office of Education, Bureau of Research, Non-Visual Perceptual Systems Laboratory, Graduate School, University of Louisville, Louisville, Kentucky, 1969.

**Dr. Emerson Foulke is Director of the Perceptual Alternatives Laboratory and E. McLean Wirth is a former research assistant at the Center for Rate-Controlled Recordings, University of Louisville, Louisville, Kentucky 40208.

A computer may also be used for the time compression of speech (Cramer, 1968; Scott, 1965). In this approach, the recorded speech signal is temporally segmented, some of the time segments are discarded according to a sampling rule for which the computer has been programmed, and the remaining segments, abutted in time, are reproduced as time-compressed speech.

In a scheme proposed by Scott (1967a), the signal resulting from the process just described is applied to one earphone of a headset. The samples that would have been discarded in the kind of compressed speech described heretofore, are retained, abutted in time, and supplied to the other earphone. With this approach, for compressions in time of 50% or less, all of the recorded signal is preserved in the compressed reproduction. It is only rearranged temporally. For compressions greater than 50%, some of the signal must be discarded, but much more is preserved than when only one succession of samples is reproduced. Scott calls the product of this process "dichotic speech."

When speech is compressed by an electromechanical compressor of the Fairbanks or Springer type, a single file of time-abutted samples is reproduced and this method will be referred to hereafter as the single file sampling method. When a computer is used to produce dichotic speech, two parallel files of time-abutted samples are reproduced, and this method will be referred to hereafter as the double file sampling method.

When speech is compressed in time by discarding samples of the original signal, as the length of samples is reduced, the probability is reduced that a critical feature of a speech signal will fall entirely within a discarded sample (Garvey, 1953b). In designing a speech compressor, the physical parameters of the system must be adjusted to produce discard samples, the durations of which are short enough so that the probability of discarding a critical feature of a speech signal can safely be ignored. Two types of speech compressors have been developed for commercial distribution. One is based directly upon the Fairbanks scheme.* The other, based directly upon the

---

*The speech compressor now manufactured by Mr. Wayne Graham, Discerned Sound, 4459 Kraft Avenue, North Hollywood, California 91602, is based upon the Fairbanks design.

Springer scheme, is the Information Rate Changer.* The Fairbanks scheme permits adjustment of the duration of discarded samples. In the Springer scheme, this capability is sacrificed in the interest of convenience of operation.** In either case, however, samples are discarded, and there is some probability that one or more of these samples may contain a critical feature of a speech signal. Since the process resulting in dichotic speech discards none of the speech signal in the range of compression bounded by zero and 50%, the probability of discarding a critical feature of a speech signal should be reduced to zero. Consequently, a reasonable conjecture would be that, in the long run, words compressed by the process resulting in dichotic speech should be somewhat more intelligible than words compressed by discarding samples of the speech signal. The superior intelligibility of dichotic speech might not be manifested on any given comparison of the two alternative reproductions of a single word. However, as the length of the list of words used for such a comparison was increased, there would be an increased opportunity for the sampling accidents that can occur with the single file sampling method, and the relative superiority of dichotic speech should begin to emerge. Accordingly, an experiment was performed in which a list of words, compressed by the two methods just described, were compared with respect to intelligibility.

Method

Subjects

Sixty Ss, of both sexes, enrolled in introductory psychology classes at the University of Louisville, served in the experiment. Subjects had no obvious hearing defects, and little or no prior experience in listening to time-compressed speech.

Apparatus and Materials

A list of 100, phonetically balanced words was read orally by a professional reader in the Talking Book Studios of the American Printing

---

*The current version on the Springer device, known as the Information Rate Changer, is distributed in this country by Infotronic Systems, Inc., 2 West 46th Street, New York, New York 10036.

**The duration of the discarded samples produced by the Information Rate Changer, a commercial device embodying the Springer scheme, is 30 milliseconds.

House for the Blind, and recorded on magnetic tape by means of an Ampex tape recorder, model 300. This "master tape" supplied the input to a speech compressor of the Springer type, constructed at the University of Louisville, and to the computer used in preparing dichotic speech.* Since the samples discarded by the electromechanical speech compressor were 40 milliseconds (msec.) in duration, the computer was adjusted so that the samples normally discarded, but retained by the computer for dichotic presentation, were 40 msec. in duration, too. The master tape was reproduced, by both methods, in 47%, 44%, 41%, 39%, and 37% of the original production time. If a recording of connected speech, occurring at the average oral reading rate of 175 words per minute (wpm) (Foulke, 1969), were subjected to these compressions, the resulting word rates would be 375, 400, 425, 450, and 475 wpm. Compressions in this range were chosen because earlier research (Garvey, 1953b; Fairbanks & Kodman, 1957; Kurtzrock, 1957) indicated that words presented at more moderate compressions would have been completely intelligible, with either kind of compression. The compressed reproductions were copied on magnetic tape for presentation in the experiment. In the case of dichotic presentation, the normally retained samples of the compressed signal were recorded on one track of a two-track stereo tape, while the normally discarded samples were recorded on the other track. Of course, only one track was required for recording the output of the electromechanical compressor. These tapes were reproduced, during the experiment, on a Revox tape recorder, model G36-III. The tape recorder was connected through a Pilot stereo preamplifier model 216A, and a Pilot stereo amplifier model SA-260 to a pair of Western Electric headphones, type ANB-H-1, equipped with ear cushions, and wired for stereophonic listening. When the tape containing speech compressed by the double file sampling method was reproduced, the file of samples recorded on one track of the tape was presented to one ear, and the file of samples recorded on the other track was presented to the other ear. When the tape containing words compressed by the single file sampling method was reproduced, the same signal was presented to both ears. The E monitored the experiment by listening to another pair of earphones, connected to an auxiliary output on the tape recorder.

Procedure

The 60 Ss were divided into five groups, with 12 Ss in each group. Each group was tested with words presented at only one of the five compressions represented in the experiment. Six members of each group heard the first 50 words in the list, compressed by the double file sampling method. The remaining 50 words were compressed by the single file sampling method. For the other six members in each group, the first 50 words in the list were compressed by the single file sampling method, while the remaining words were compressed by the double file sampling method and presented as dichotic speech. This precaution was taken to control for the possibility that some words may have been treated more favorably by one method or the other. To control for the possibility of an effect due to order, three of the Ss in each sub-group heard words compressed by the double file sampling method, followed by words compressed by the single file sampling method. The order of presentation was reversed for the remaining three Ss in each sub-group.

---

*Dichotic speech was prepared for this experiment at the National Security Agency, Fort George G. Meade, Maryland, by John Boehn, using methods developed by Dr. Robert Scott. Dr. Scott's assistance in arranging for the preparation of this material is sincerely appreciated.

Subjects were tested one at a time. Each S wrote the words he thought he heard on an answer sheet in numbered answer spaces. Approximately five seconds elapsed between the onsets of consecutive words. Subjects were instructed to guess if they were uncertain about a word.

Results

At each fraction of original production time represented in the experiment, two scores were determined for each S--the number of words compressed by double file sampling that were missed, and the number of words compressed by single file sampling that were missed. Means and standard deviations of error scores are shown in Table 6.1. The influence of the method of compression upon the relationship between the amount of compression and error frequency is graphed in Figure 6.1. In this figure, the fraction of original production time required for compressed reproduction, at each of the five compressions represented in the experiment, is scaled on the x-axis. Fractions are expressed as percents. The entry recorded below each scaled value on the x-axis is the word rate that would result if a listening selection, read at the average oral reading rate of 175 wpm, were reproduced in the fraction of original production time indicated by that value. The y-axis is scaled in terms of error scores. This figure indicates an orderly growth in error scores as the fraction of original production time required for compressed reproduction is reduced. On the other hand, the differences associated with the methods of compression appear to be small and unsystematic.

The apparent outcome of the experiment was checked by an analysis of variance of error scores, with scores classified according to amount of compression and method of compression, and with repeated measures on the methods variable. The results of this analysis are shown in

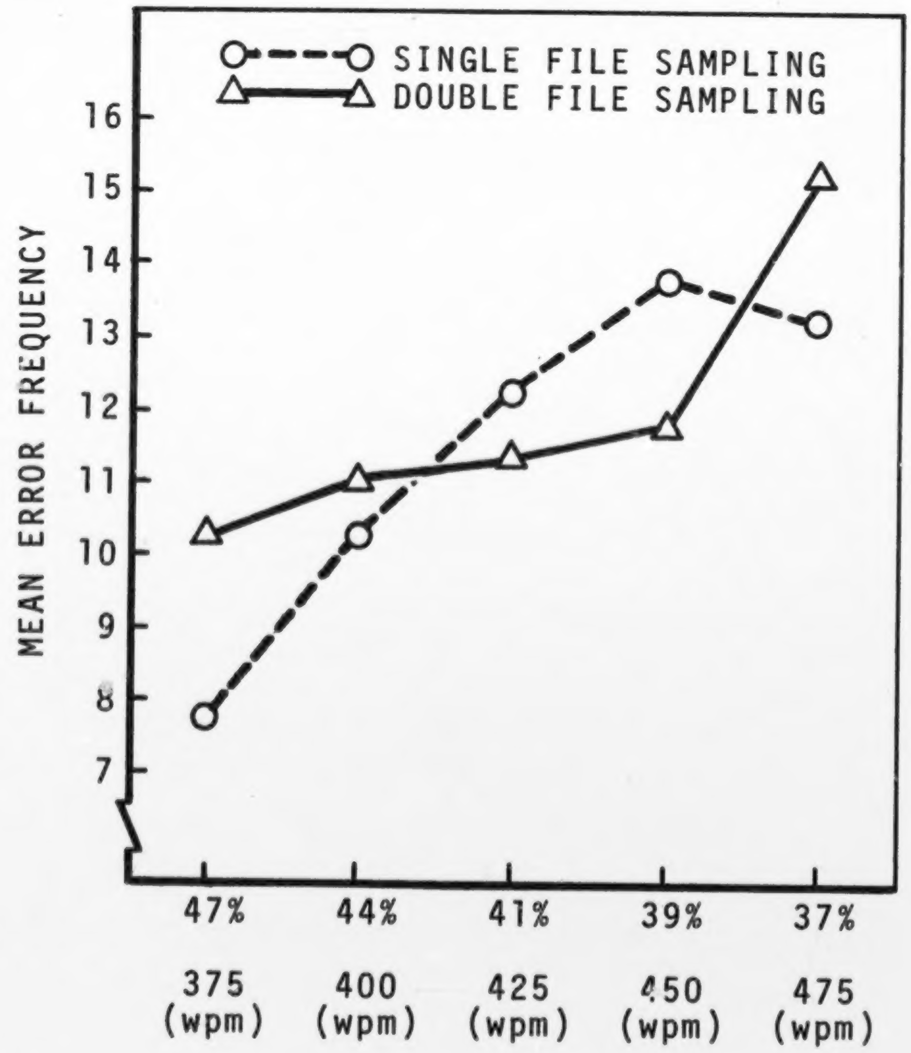


Figure 6.1. Identification errors as a function of compression in time with method of compression as the parameter.

Table 6.2. The growth in errors accompanying the reduction of time available for compressed reproduction was significant at the .01 level, but the variance associated with the method of compression did not reach significance at the .05 level. The interaction between these variables was significant at the .05 level.

TABLE 6.1

IDENTIFICATION ERRORS FOR WORDS COMPRESSED BY SINGLE AND DOUBLE FILE SAMPLING

Percent of Original Production Time	Method of Compression			
	Single File Sampling		Double File Sampling	
	Mean # of Errors	SD	Mean # of Errors	SD
47%	7.92	2.80	10.25	2.81
44%	10.25	3.59	10.92	3.97
41%	12.33	2.53	11.25	3.63
39%	13.83	4.08	11.75	3.00
37%	13.25	4.17	15.17	3.34

TABLE 6.2

ANALYSIS OF VARIANCE OF IDENTIFICATION ERRORS

Source of Variation	df	MS	F
Level of Compression	4	93.51	5.26**
Error (between)	55	17.76	
Method of Compression	1	3.68	0.46
Level X Method of Compression	4	21.70	2.71*
Error (within)	55	8.00	

*p < .05
**p < .01

A test of simple main effects was made in order to examine the influence of method more closely. The results of this analysis are shown in Table 6.3. The significant fact recorded in this table is that differences in error scores as a consequence of the method of compression used were not significant except for those words compressed to 47% of original production time.

TABLE 6.3

ANALYSIS OF VARIANCE OF SIMPLE MAIN EFFECTS

Source of Variation	df	MS	F
Method of Compression for 375 wpm	1	32.67	4.08*
Method of Compression for 400 wpm	1	2.67	0.33
Method of Compression for 425 wpm	1	7.04	0.88
Method of Compression for 450 wpm	1	26.04	3.25
Method of Compression for 475 wpm	1	22.04	2.75
Error	55	8.00	

*p < .05

The Newman-Keuls Test for Ordered Pairs of Means was performed in order to determine the effect of compression more precisely. Since differences due to method were, with one exception, not significant, the error scores obtained at each fraction of original production time were pooled. The results of this analysis are shown in Table 6.4. This table is arranged in matrix form, with the fractions of original production time in which words were reproduced displayed in decreasing order along the top, and down the left hand margin of the table. Entered in each row, under the appropriate column headings, are the fractions of original production time for which error scores were not significantly different from the error score associated with the fraction of original production time, recorded in the left hand margin, which identifies that row. If the table is examined as a whole, the effect of the compression variable is depicted by the total array of entries in the table.

TABLE 6.4

NEWMAN-KEULS TEST FOR ORDERED PAIRS OF MEANS

Fraction of Original Production Time	47%	44%	41%	39%	37%
47%	47%	44%	41%		
44%		44%	41%	39%	
41%			41%	39%	37%
39%				39%	37%
37%					37%

Discussion

A significant interaction between method and amount of compression would be an interesting finding. However, since the general effect of varying the method of compression was not statistically significant, and since the differences at the various fractions of original production time were unsystematic and insignificant with one exception, the interaction that was found in the present experiment is probably without experimental significance. Where it was observed, the difference in favor of dichotic speech was probably the accidental result of uncontrolled factors in the experiment, such as differences in the recording quality of the tape bearing the words used in this comparison, or a higher frequency of sampling accidents in the 50 words processed by the electromechanical compressor.

The intelligibility of words compressed by double file sampling has been compared with the intelligibility of words compressed by single file sampling in an experiment reported by Gerber (1968). His results cannot be directly compared with the results of the present experiment, since the words he used for testing were reproduced in 50% of original production time or more, while the words used in the present experiment were reproduced in less than 50% of original production time. In Gerber's experiment, words were compressed to 75%, 67%, and 50% of original production time and, at each compression, samples with durations of 30, 40, and 50 msec. were discarded. In all of the nine comparisons provided by his experiment, he found a difference in favor of dichotic presentation. When the discarded samples were 50 msec. in duration, this difference was significant at all three compressions. However, in the six comparisons in which the discarded samples were 30 and 40 msec. in duration, three of the differences were statistically insignificant, and the remaining three, though significant, were relatively small.

The fact that Gerber found a consistent difference in favor of dichotic presentation, when the discarded samples were 30 and 40 msec. in duration, while the present experiment revealed no consistent advantage for dichotic presentation, may be, in part, a consequence of differences in the range of the compression variable explored by the two experiments. Since, in Gerber's experiment, none of the words were reproduced in less than 50% of original production time, dichotic presentation preserved all of the original speech signal. Since, in the present experiment, all the words were reproduced in less than 50% of original production time, dichotic presentation did not completely eliminate the necessity of discarding some of the speech signal. Even though discarded samples are quite small when double file sampling and dichotic presentation are used to reproduce words in less than 50% of original production time, sampling accidents are still possible, and may have injured the intelligibility of some of the words that were presented dichotically in the present experiment.

Though Gerber feels that his experiment has demonstrated the superiority of dichotic presentation, it seems to this writer that the differences he found, even when statistically significant, were too small to be of practical significance, except when the discarded samples were 50 msec. in duration. Of course, when speech is compressed by single file sampling, and when discarded samples are 50 msec. in duration, it is probable that some of the critical features of speech signals will fall entirely within discarded samples. If single file sampling is to be successful, the discarded samples must be kept short enough so that every critical feature of a speech signal has the opportunity to be sampled. As Garvey has shown (1953b), this condition is met fairly well when the discarded samples are no longer than 40 msec. in duration. In general, it can be said that the intelligibility of words is preserved better by double file sampling than by single file sampling when the discarded samples are long enough so that some of the critical features of speech signals can fall entirely within discarded samples, but that as the duration of discarded samples is shortened, the superiority of double file sampling is diminished. The results of both Gerber's experiment and the present experiment suggest that at 40 msec., this superiority has nearly vanished. Though the experience of listeners, and the examination of spectrographic records (Foulke, 1969), suggests that critical features of the speech signal may occasionally be insufficiently sampled when the discarded samples are 40 msec. in duration, the effects of such sampling accidents are counteracted by other factors, such as the listener's knowledge of the sequential dependencies inherent in sequences of phonemes and syllables.

CHAPTER VII

THE INTELLIGIBILITY OF COMPRESSED WORDS

Robert Heise*

A consistent finding in the literature on compressed speech at the time of the first Louisville Conference on Time-Compressed Speech in 1966 was that speech comprehension had been shown by many authors to be relatively accurate until the speech rate was more than doubled. Beyond 50% compression, comprehension suffered severely.

Another consistent finding at that time was that single word intelligibility withstood the ravages of compression, by the sampling method, much better than connected discourse. It was commonly found on intelligibility tests involving word lists that single words could be recognized accurately at speech rates well beyond the rate where the comprehension of connected discourse declined.

Foulke and Sticht (1967b) advanced a tentative explanation for this inconsistency. Their two process hypothesis contended that the process of speech comprehension entails the registration, encoding, and storage of information, and that these operations require time. The implication was that the comprehension of connected discourse declined beyond a critical speech rate because a listener's capacity to perform the necessary analytical operations was surpassed. Since word intelligibility had been shown to decline at a different speech rate than comprehension, the authors suspected that single word intelligibility was not the critical factor.

Foulke (1968b) went on to provide further evidence that variations in single word intelligibility exerted little or no influence on the comprehension of connected discourse. He varied the vocal pitch of connected discourse while holding speech rate constant. Although shifts in the vocal pitch of speech had been shown by Garvey (1953b) to be very detrimental to word intelligibility, comprehension was unaffected.

---

*Mr. Robert Heise is a graduate research assistant in the Perceptual Alternatives Laboratory, University of Louisville, Louisville, Kentucky 40208.

Another experiment by Foulke and Sticht (1966), in which they varied the original speaking rate, then compressed the passages by three different amounts, again holding the rate of presentation constant, showed that speech comprehension was not measureably influenced. Even though the duration of the individual words must have been different for the three passages, comprehension was unaffected.

The two experiments just cited reflect a significant departure from past inquiry into the relation of word intelligibility and speech comprehension. The earliest evidence on this relation was obtained by testing for intelligibility with phonetically balanced lists of monosyllables without concern for the fact that the words in the lists bore an unknown likeness to the words occurring in the listening tests for comprehension. Thus, the innovation in the experiments by Foulke (1968b) and Foulke and Sticht (1966), was that they varied the intelligibility of the actual words that occurred during the passages used to test comprehension.

In the summer of 1968, Emerson Foulke and I confronted the issue of the relation between intelligibility and comprehension with what must be called a new slant to an old tactic (Foulke & Sticht, 1967a). We measured intelligibility by means of a word list but included in this list only words that occurred in the passages used to test comprehension. This was accomplished by sampling a representative amount of words from all of the different words that occurred in the passages.

The test used to measure listening comprehension in this experiment was the Nelson-Denny Reading Test which consists of eight selections of general interest with questions after each. The selections were recorded with a male speaker. Six versions were prepared for testing, one at normal speed and five compressed versions. Compression was accomplished by the sampling method.

The intelligibility test consisted of 159 words of from one to five syllables in length. The words were chosen to represent all the possible beginning sounds of words in the language. These words were spoken within a surrounding carrier phrase so that the reader could modulate his voice evenly for all words, and so that the contour of the words would closely approximate their pronunciation in the listening selections.

Twelve groups, of 10 college Ss each, were necessary. Six groups had the intelligibility test before the comprehension test and the other six had the comprehension test first.

The results of the testing of comprehension (Figure 7.1) showed that the scores dropped only 6% in the range where 50% or more of the original signal remained, but that comprehension declined severely thereafter (50% compression represented about 325 words per minute [wpm]).

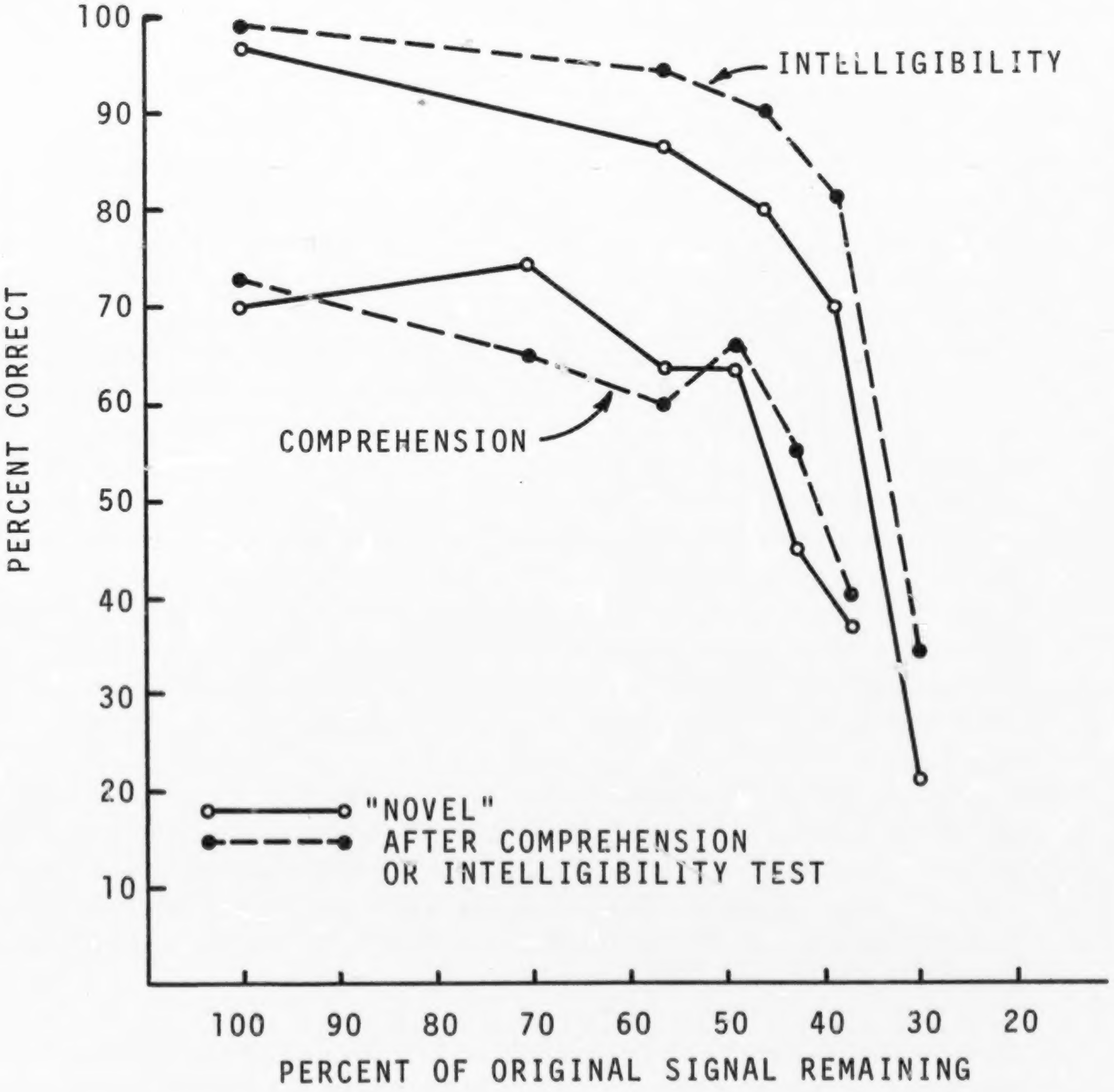


Figure 7.1. Listening comprehension and word intelligibility as a function of compression.

Intelligibility scores remained above comprehension scores at all measured points, showing the most severe drop beyond 60% compression. These results were quite straightforward.

However, it was of special interest to find that when the comprehension test preceded the intelligibility test, the effect was to enhance single word intelligibility consistently by about 10% at all the five compressed rates. Apparently, hearing the words (without listening for them) during the comprehension test provided the Ss with some context for recognition during the intelligibility testing.

Prior to this finding, it was known that a number of authors studying the intelligibility of compressed words familiarized their Ss with the words in the lists used for testing, and, as a consequence, obtained higher intelligibility scores in a range of compressions comparable to those used in this experiment. So, we decided to provide some familiarization with the words in a second experiment.

In the second experiment, the same list of words used in the first experiment was presented once at a normal rate prior to each of the five compressed presentations. It was of further interest to compare the effect of the carrier phrase in the first experiment. This was accomplished by presenting the list of words with each word in isolation, that is, without any surrounding linguistic context.

The results proved interesting for two reasons: (1) one prior presentation of the list before testing did enhance scores by 10% on the average, and (2) when the intelligibility scores for the first experiment were compared, recognition was 4% higher for the words heard in isolation. This second result can be seen in Figure 7. 2.

Concerning the first finding, it is not suggested that the effect of hearing the words during the passages provided the same kind of "familiarization" that was provided when the complete list was heard. The Ss were evidently able to use the meaningful context of the words when the test followed the passages, but what was gained when the list of words was presented, per se, was an explicit limiting of the alternatives to which the listeners had to direct their attention.

The second finding, presented in Figure 7. 2, was unexpected. We supposed that the Ss could derive temporal cues from the words preceding the test words when they were embedded in the carrier phrase, "You will write (the test word) now." These cues would be completely lacking in the case of the words presented in isolation, and we expected scores to be lower in this case. Very simply, since the words were the same in both cases, we supposed that the carrier would facilitate recognition.

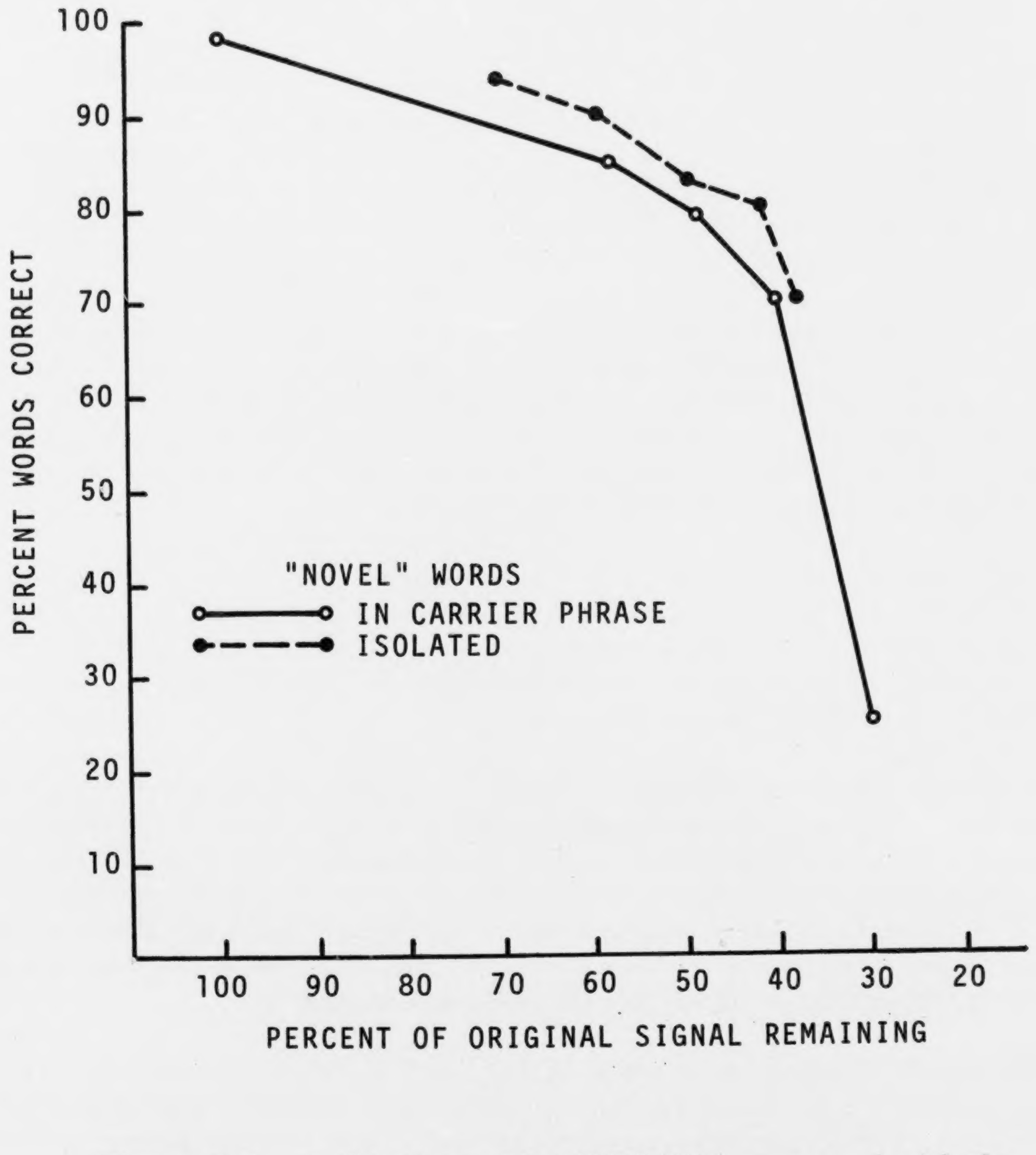


Figure 7.2. Isolated word intelligibility compared with the intelligibility of words presented in a carrier phrase.

We discovered, however, when we listened to the words ourselves and compared the duration of the words pronounced by our reader in the carrier phrase to the same words pronounced by him in isolation, that the isolated words were greater in duration. It came to our attention after the experiment that this phenomenon had been commented on before in the literature. When words are pronounced without any surrounding linguistic context, they tend to be elongated. We felt that this was the explanation for the consistently higher scores, notwithstanding the supposed temporal value of the carrier phrase.

Since the results indicated that it was the duration of the individual words that influenced their recognition, and also that familiarity with them was in another sense a factor, we decided to explore the possibility that words encountered more frequently in daily usage and words of more syllables would be more intelligible. Rather than implement another experiment, we analyzed the written responses of our Ss on the first experiment.

The 159 words were categorized in terms of the number of syllables in each, and their Thorndike-Lorge frequency of occurrence in the American language. Since this analysis was not planned and every word has the characteristics of length and familiarity, we had to be content with a relatively small number of words where the effect of these two factors could be observed independently.

Table 7.1 describes the 159 words in terms of length and familiarity. This table shows that the majority of the most familiar words (in the top row) were of one and two syllables in length, and that the least familiar words (the bottom row) tended to be of more syllables. Words of an intermediate frequency of usage also fit the pattern. In short, lengthy words tended to be unfamiliar.

The combinations of the characteristics of length and familiarity probably account for the results shown in Figure 7.3. This figure shows that word length had a different effect on intelligibility for lower speech rates than for higher. With reference to the upper three curves, it can be seen that where more than 40% of the original signal remained, word length exerted little effect on word recognition. However, for the two lower curves, there was a consistent and marked decline in the intelligibility of longer words.

It was of considerable interest to identify the effects of these two factors independently. I will speak only of the most familiar and least familiar one and two syllable words.

It can be seen in Figure 7.4 that more familiar words (these are the upper curves) were always more intelligible, and that this tendency was magnified somewhat as more of the signal was discarded. (The mean

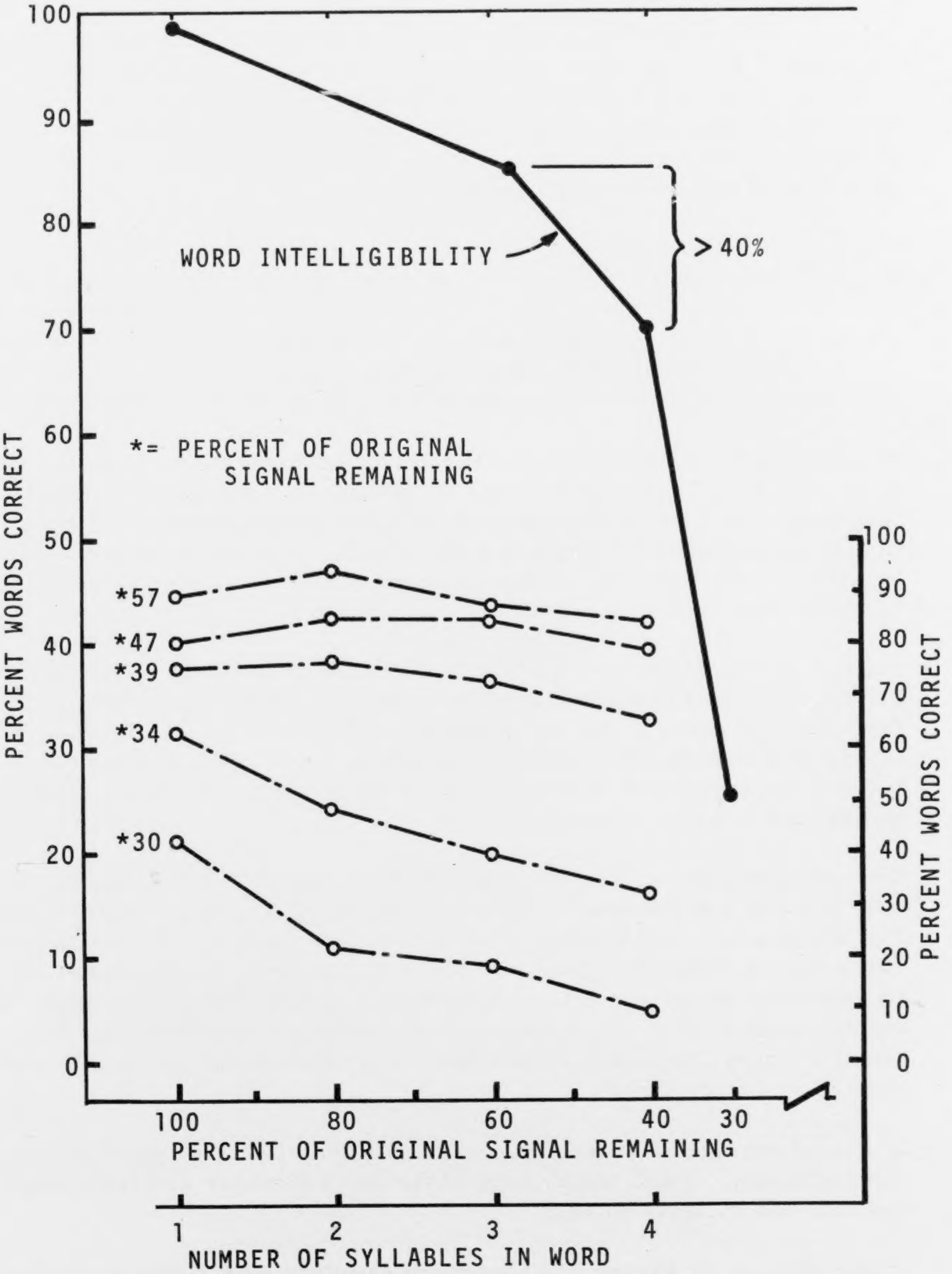


Figure 7.3. Word intelligibility as a function of compression and syllable intelligibility at five different amounts of compression.

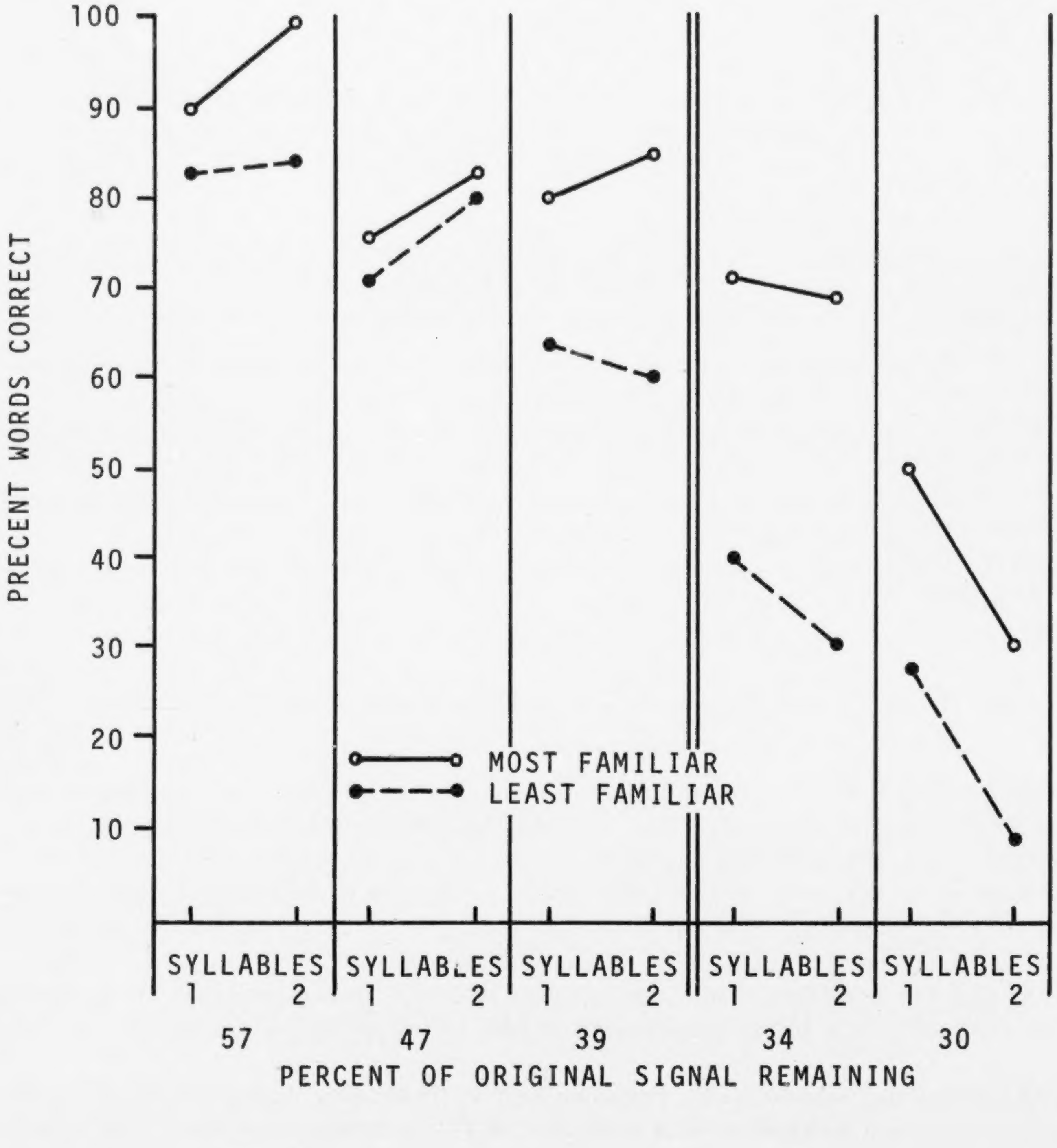


Figure 7.4. Word intelligibility as a function of compression. The parameters are word length in syllables and frequency of usage.

differences between the upper and lower curves, from left to right, are about 10%, 4%, 20%, 37%, and 20%.) Also, as the speech rate was increased, the effect of an additional syllable was to enhance word intelligibility where 40% or more of the original signal was present, but to depress intelligibility thereafter. This tendency was most clear in the case of the most familiar words, and less so for the least familiar.

TABLE 7.1

A DESCRIPTION OF THE WORD LIST IN TERMS OF THE NUMBER OF SYLLABLES IN EACH WORD AND THORNDIKE-LORGE FREQUENCY

Thorndike-Lorge Frequency	Number of Syllables in Word					
	1	2	3	4	5	
500 most common	34	13	1			
100 or more per million	9	14	3			
50 or more per million	8	9	1	3		
Less than 50 per million	8	23	23	9	1	
Totals	59	59	28	12	1	159

The finding that lengthier words were less intelligible at the highest speech rates seems to contradict the finding that words of greater duration were more intelligible in Experiment II. This finding calls for an additional factor as a variable in the perception of single words under compressed speech conditions. We suggest that this factor is the phonetic structure of words of many syllables. Within this suggestion is the implication that the recognition of a word requires a molecular level of analysis of its content. We found consistent evidence to fulfill this requirement.

All 159 of the words were transcribed syllabically and phonetically, and every written response was matched with its transcription. The results of this endeavor showed that the average number of words rendered correctly was 61.38%, the average number of syllables correct was 65%, and the average number of phonemes correct was 76.6%. Thus, although a listener missed a word, it cannot be said that he missed it completely. In fact, about 15% more information in phonemic terms is perceived than a higher level of analysis would indicate, that is, the level of the whole word.

CHAPTER VIII

STUDIES ON THE EFFICIENCY OF LEARNING BY LISTENING TO TIME-COMPRESSED SPEECH

Thomas G. Sticht*

One of the intriguing aspects of the use of time-compressed speech is that more information can be presented in a given amount of time. For instance, if a message is time-compressed by 50%, it is possible to present the compressed version two times in the same amount of time required to present the uncompressed version once. A second alternative is that extra information may be presented in the time saved by the compression process.

Both of these possibilities were obvious to Fairbanks, Guttman, and Miron (1957a, b) in their work which introduced the automated time compression process to the experimental investigation of learning by listening. In one of their studies they compared the comprehension of material compressed by 50% (282 words per minute [wpm]), but presented twice, with the comprehension of the identical, uncompressed material requiring the same amount of time for presentation. They found a slight increase in comprehension with the repeated time-compressed message over that obtained with the single presentation of the uncompressed message. Their work also indicated that the double presentation appeared slightly more successful with men of moderate than of high mental aptitude.

The work of Fairbanks and his associates suggested to me that the repetition procedure might prove even more successful with very low aptitude men. I was also interested in finding out if the comprehension of repeated time-compressed messages might be different for different combinations of compression. For instance, a message compressed by 40% and repeated at a compression ratio of 60% might produce a higher level of comprehension than if the reverse sequence was used, i.e., if the 60% compressed version was presented before the 40% compressed version. This might be so because more information could be stored from a less compressed and, hence, more slowly presented message to facilitate the comprehension of a more rapidly presented message. To check these ideas out, an experiment was performed in which the comprehension of repeated time-compressed messages, presented in several repetition sequences, was compared with the comprehension of messages presented one time in either compressed or uncompressed versions.

---

*Dr. Thomas G. Sticht is with the Human Resources Research Organization at the George Washington University, Presidio of Monterey, California 93940.

Comprehension of Repeated Time-Compressed Messages

For this study, a selection on the use of Carbon 14 for dating relics, taken from Form 1A of the Sequential Tests of Educational Progress, was used. The tape-recorded listening selection was compressed by an Eltro Information Rate Changer to produce compression ratios of 0, 36, 46, 53, and 59%. In wpm rates, these compression ratios correspond to normal (175), 275, 325, 375, and 425 wpm. A 20-item "fill-in-the-blank" test was prepared to evaluate listening comprehension.

These listening selections were grouped to form four pairs of repeated messages. One tape presented the passage compressed by 36% and repeated at 59%. The remaining tapes were paired to produce compressed message sequences of 59% followed by 36%, 46% followed by 53%, and 53% followed by 46%. When paired in this way, the combinations of 36% and 59% required 105% of the time required to listen once to the uncompressed message, and the combination of 46% and 53% required 101% of the normal listening time.

Subjects of high and low mental aptitude were selected from Army inductees who scored high or low on the Armed Forces Qualification Test (AFQT). In terms of intelligence, these groups represented men having IQs of around 120 plus, and 90 or below (Hedlund, 1969). Individual Ss listened to the tapes in a sound-deadened room. They listened to both levels of compression of the repeated message before taking a 20-item "fill-in-the-blank" comprehension test which was presented aurally.

The results of the experiment are summarized in Figure 8.1. This figure compares the comprehension of the repeated messages with the comprehension of the same listening selection when presented one time at compression ratios of 0%, 36%, and 59%. These data were obtained from Ss tested in previous research (Sticht, 1968) who were matched with the present Ss on AFQT.

As Figure 8.1 indicates, both high and low aptitude Ss showed improved comprehension with the repeated selections. However, in no case did the double presentation improve comprehension over that obtained with a single presentation of the uncompressed selection. The only suggestion

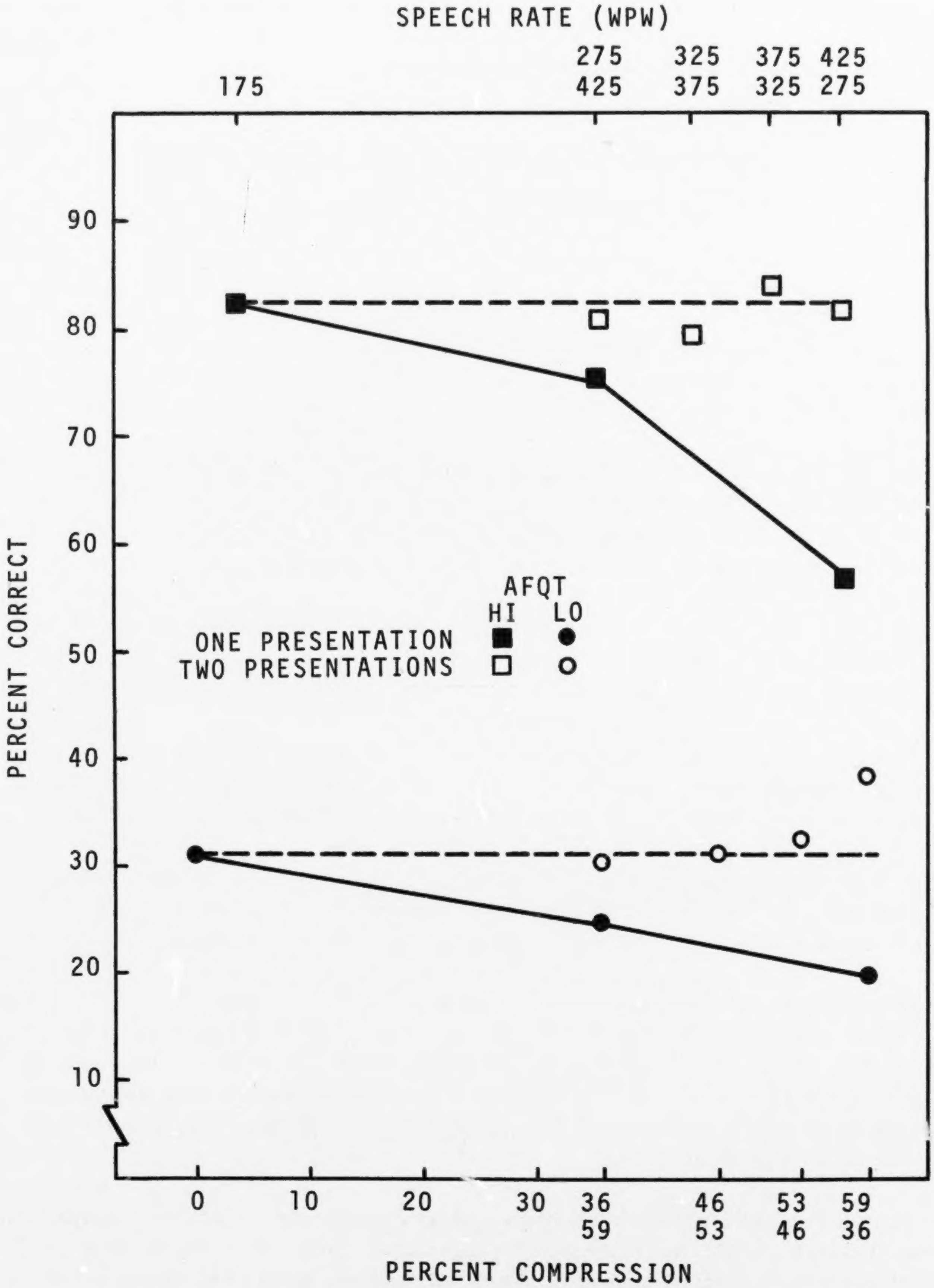


Figure 8.1. Comprehension scores for high and low aptitude Ss for single (filled symbols) and double presentation of listening selection (see text for explanation of abscissa).

that performance may have improved over that for a single presentation is with the lower mental aptitude group with the 59%-36% sequence. However, the difference was not statistically significant (Fischer exact probability test).

A feature of note in Figure 8.1 is that, in both double presentations at 53% and 46% compression, comprehension was better than with a single presentation at either of these compression ratios. Many studies (Foulke & Sticht, 1969) have reported a notable decrease in comprehension with single presentations of messages at word rates of 325 or 375 wpm. The present results indicate that, for both ability groups, some savings occurred from listening first at either 375 or 325 wpm before listening to a repetition of the message at these word rates (in this regard see Jester & Travers, 1967). But apparently the savings was not sufficient to raise the performance level above that for the single presentation of the uncompressed passage.

As mentioned earlier, Fairbanks et al. (1957b) obtained results similar to those of the present research. A double presentation of materials compressed by 50% (282 wpm) resulted in a very slight improvement in comprehension over that for a single presentation of the uncompressed message. Their results, the work of Friedman, Graae, and Orr (1967), Hopkins (1969), and the present results seem to indicate that using the extra time resulting from the time compression of materials to simply repeat information is not likely to improve learning over what could be obtained by listening once to the uncompressed message presented within the "normal" range of speech rate. Furthermore, the work of Fairbanks et al. (1957b) suggests that listening twice to the uncompressed message is not likely to produce very drastic improvements in comprehension--if any at all. Possibly the effectiveness of repeated time-compressed messages may be increased for Ss who are trained in listening to time-compressed speech, but there is no firm data to suggest this (cf., Friedman et al., 1967).

On Learning More Per Unit of Time by Means of Time-Compressed Speech

A second possibility which has been mentioned for improving the efficiency of learning by listening is to use the time saved by the compression of material to present additional information. Fairbanks et al. (1957a) used the time saved resulting from 30% (201 wpm) compression of a message to emphasize certain portions of the message. As they pointed out, this amounts to trading temporal redundancy for verbal redundancy. Their results indicated that the reinforcing of certain parts of the selection did, indeed, increase the comprehension of the emphasized materials.

However, this increase appeared to occur at the expense of the remaining unemphasized content, for the comprehension of this material showed a highly significant decline. Thus, the overall comprehension score for the reinforced compressed material was less than the overall score for the uncompressed material.

Fairbanks and his associates suggested that emphasizing certain parts of the message may have led the Ss to assume that verbal redundancy meant "important to learn" and, hence, such emphasis may have selectively focused attention upon certain parts of the message, while diminishing attention to the remainder of the material. This suggested to me that if the time saved by the compression process was used to present additional new information, perhaps an overall increase in the amount learned in a given unit of time might occur.

To evaluate this idea, research was performed in which independent groups (N = 15 per group) of high (AFQT $\geq$ 80) and low (AFQT $\leq$ 30) aptitude men listened to a recorded message presented under five different conditions. Under one condition, the men listened to the message presented at a normal uncompressed speech rate of 178 wpm. The time required to listen to the uncompressed message was 6 minutes 4 seconds. By means of the time sampling compression method, two additional versions of the message were presented. One was compressed by 36%, which produced a speech rate of 278 wpm and reduced the listening time from 6 minutes 4 seconds to 3 minutes 53 seconds. The third version of the message was produced by compressing the message by 53%. This resulted in a speech rate of 378 wpm and reduced the listening time from 6 minutes 4 seconds to 2 minutes 52 seconds. Thus, three versions of a message were available having speech rates of 178, 278, and 378 wpm and for which the time needed to listen to the message decreased from 6 minutes 4 seconds in the case of normal speech to 2 minutes 52 seconds using speech of 378 wpm. These tapes were used to assess the effects of increasing the speech rate upon the comprehension of a recorded message.

Two additional groups of high and low aptitude men listened to the test message at 278 (N = 14) or 378 wpm and then listened to additional information until their total listening time was 6 minutes 4 seconds, i.e., the same amount of time as required to listen to the normal uncompressed message. These groups thus listened to what the previous three groups had heard, plus additional information. For all conditions, Ss were assigned to the various treatment groups in an unsystematic manner, as they became available, until all treatment cells were filled.

The message used in this study was the "Roland" selection from the standardized listening passages prepared by Clark and Woodcock (1967). Subjects listened to this selection in an open classroom. They were seated in a semicircle around a tape recorder adjusted to a "comfortable" listening level determined by the group. Subjects listened first to the recorded instructions on the standardized listening tapes (they were instructed to ignore the references to earphones in the instructions). Then they listened to the "Roland" selection. Immediately after the presentation of the listening selection, Form A of the comprehension tests was administered both by reading and listening. This test contains 28 four-alternative multiple-choice questions. In the present research, the 6 minutes 4 seconds of listening time presented at a normal (178 wpm) speech rate provided answers relevant to only the first 14 of the 28 test items. This was true also for the two compressed versions in which the listening times were 3 minutes 53 seconds and 2 minutes 52 seconds. For the compressed versions in which the listening time was held constant at 6 minutes 4 seconds, information relevant to both the first and second halves of the comprehension test was presented. In this case, more relevant test information was presented in 6 minutes 4 seconds with the speech rate at 378 wpm than at 278 wpm. Of primary interest was whether or not holding the listening time of the compressed message equal to that of the uncompressed message would result in an overall increase in scores on the total 28-item test.

The results of the study are summarized in Figure 8.2. In this figure, the unfilled symbols designate the conditions for which the listening time was constant at 6 minutes 4 seconds. The filled symbols indicate those conditions for which listening time was reduced. The square symbols are for the high aptitude Ss and the round symbols are for the low aptitude Ss. The abscissa indicates the rate of speech at which the message was presented, and the ordinate is the percent correct on the 28-item comprehension test.

The data indicate that, under those conditions in which the speech rate was increased and the listening time was reduced, comprehension decreased for both high and low aptitude Ss. This is the typical finding regarding the relationship between speech compression and comprehension (Foulke & Sticht, 1969).

The data of primary interest are given by the unfilled symbols. In this case the listening time was constant at 6 minutes 4 seconds, while the speech rate was increased from 178 to 278 to 378 wpm. Thus more information was presented with the faster rates of speech. The data of Figure 8.2 indicate that, for higher aptitude men, there was no increase in comprehension scores when more information relevant to the test was presented at accelerated speech rates. For lower aptitude men, there is a suggestion that listening to additional information at accelerated speech rates may have improved comprehension over that obtained by listening to less information at the same accelerated speech rates. However, the differences indicated in Figure 8.2 are not statistically significant (t tests; p = .05).

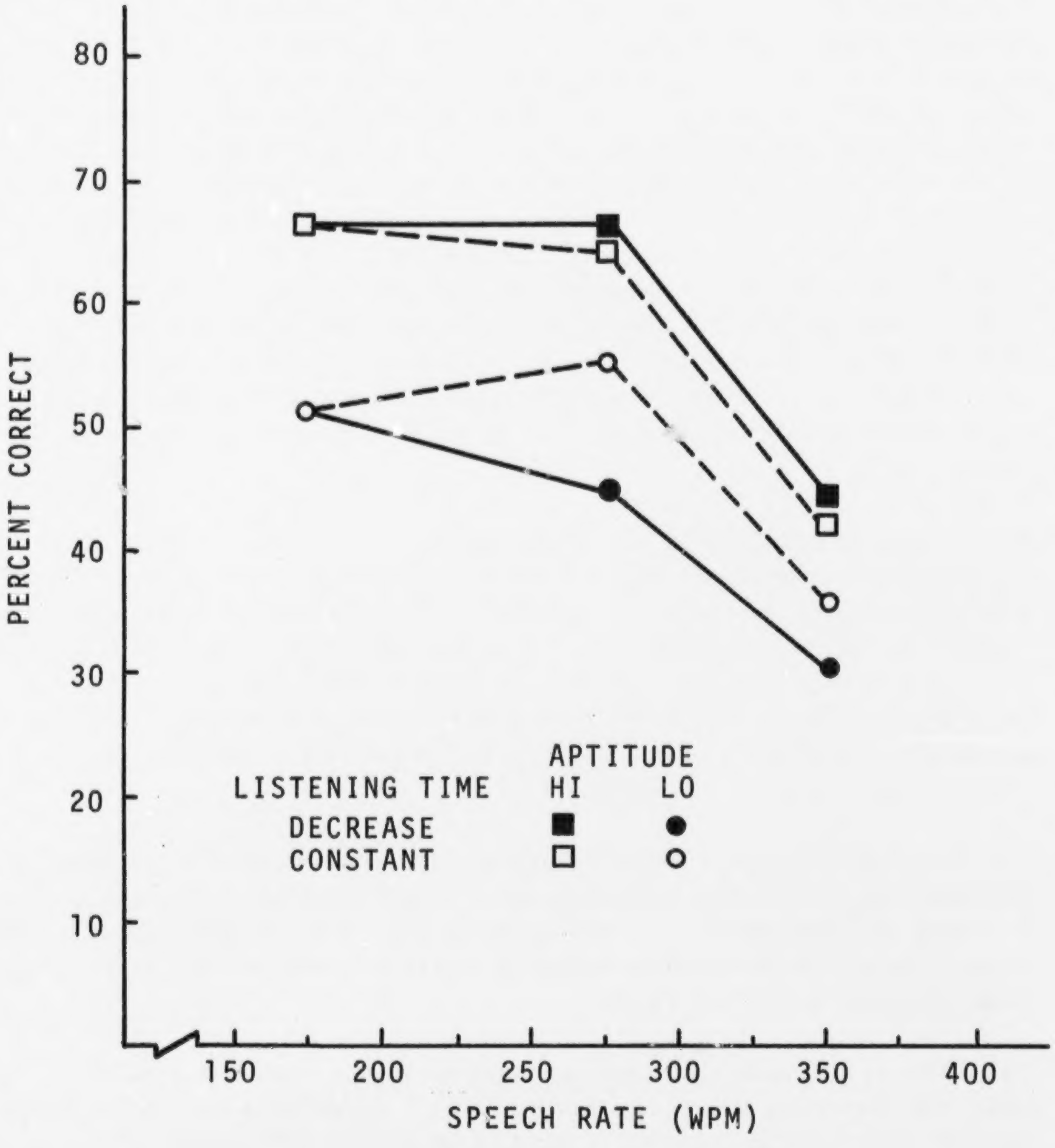


Figure 8. 2. Listening comprehension scores for high and low mental aptitude Ss who listened to time-compressed:time-limited (filled symbols) or time-compressed:time-extended selections.

These data indicate that listening to additional information in the time saved by the time compression of speech may not lead to increased learning. In the present study this was true even for material compressed only 36% and presented at 278 wpm, and even though this compression resulted in a higher "listening efficiency" score, i.e., more was learned per time spent listening, than obtained with the normal (178 wpm) rate of speech. This was so for both aptitude groups.

As mentioned earlier, Fairbanks et al. (1957a) attempted to increase learning by using the extra time resulting from 30% compression to emphasize certain content in a recorded message. They found that, whereas the learning of the emphasized materials was, indeed, improved, the learning of the unemphasized materials declined, and the total comprehension score stayed about the same as that obtained by listening to the message presented at a normal rate of speech and without added emphasis. Fairbanks et al. (1957a) suggested that emphasizing certain content might have caused Ss to consider that content as more important than the remaining content, and hence they may have ignored the unemphasized materials. They also mentioned the possibility that the response to the emphasized version of the message may have actively inhibited the response to the unemphasized content.

The present results are essentially the same as those found by Fairbanks and colleagues. But in the present case, the possibility of selectively focusing attention through emphasis of materials was avoided and, hence, does not explain the failure to find improved learning with extended listening. However, the notion of inhibition may be related to the present findings. An analysis of the responses of high aptitude Ss to the first and second halves of the 28-item test indicated that, with the materials presented at 278 wpm, the scores on the first half of the test decreased slightly when the message was 6 minutes 4 seconds in duration as opposed to when the message length was only 3 minutes 53 seconds in duration. Thus, there is the possibility that retroactive inhibition may have occurred such that listening to additional material may have interfered with the retention of previously presented material. However, the evidence for this is very slight. Also, this interpretation is not confirmed by the data for the low aptitude men who, in fact, showed a slight increase in performance for both halves of the test when listening time at 278 wpm was extended from 3 minutes 53 seconds to 6 minutes 4 seconds.

Comments on the Efficiency of Learning by Listening to Time-Compressed Speech

To recapitulate briefly, several research studies have attempted to demonstrate that the time saved by time compression might be used to increase learning over that which could be obtained by listening once to the uncompressed materials. These studies have used the time saved by the compression process to repeat or review messages (Fairbanks et al., 1957b; Friedman et al., 1967; Hopkins, 1969; the present work), to selectively elaborate parts of messages (Fairbanks et al., 1957a), or to present additional new, but related, information (the present research). To date, none of these techniques have been found to significantly increase the amount of learning over that obtained from a single presentation of the same, or unelaborated, or less extensive material presented in an uncompressed format with speech rates between 140-178 wpm.

On the basis of these limited data it appears as though the technique of trading time for information has not resulted in more information being processed by the listener for short-term retention. Most significantly, this has been true for materials compressed to speech rates of 275-300 wpm for which listening "efficiency," i.e., the amount learned per unit of listening time, has actually been higher than obtained with "normal" materials. Thus, the implication that, because of improved listening efficiency, more information can be learned in a unit of time with moderate compression has yet to be substantiated.

I would mention, however, that there are several features of the various research efforts under consideration which serve to limit conclusions to be drawn from them. For one thing, practically all of this research has involved listeners untrained in listening to compressed speech. This, of course, requires no further comment. Secondly, all of the studies have used materials within a given subject matter area. Possibly the probability of interference factors might be reduced if a different type of content was presented in the time saved by the compression process. Thirdly, these studies have presented the additional information in a single sitting and immediately tested for learning. Perhaps some spacing of the presentation of new compressed information might increase learning over that obtained by continuously listening for the same amount of time to uncompressed materials (but see Friedman et al. [1967] for preliminary negative findings using long-term intervals between repetitions of materials).

As a final comment upon the efficiency of learning from moderately time-compressed speech, it should be pointed out that the studies reported in this paper have all been concerned with using the time saved by the time compression process for increasing the learning of a given group of Ss. An alternative would be to use this timesavings for other purposes, such as instructing additional students. Thus, the efficiency of time-compressed listening does not rest solely on demonstrating an increase in the amount of learning per group per unit of time, but also by the demonstration that more groups per unit of time can be instructed with moderate compression. This is fait accompli in the many studies, including the present ones, which demonstrate that much learning can occur with materials that have been compressed by 30-40%. Clearly this timesavings can be used to instruct additional listeners. It is also obvious, and I'm sure this fact has not escaped many of you who have doggedly tested several groups of Ss in a single day, that the time saved by the compression process can also be used by researchers to recover between treatment groups.

ACKNOWLEDGMENTS

The research reported in this paper was performed at HumRRO Division No. 3, Monterey, California, under Department of the Army contract with the Human Resources Research Organization; the contents of this paper do not necessarily reflect official opinions of policies of the Department of the Army.

Reproduction in whole or in part is permitted for any purpose of the Department of the Army.

CHAPTER IX

THE APPLICATION OF RATE-CONTROLLED RECORDINGS IN THE CLASSROOM

Richard W. Woodcock*

Several interrelated studies pertaining to the application of rate-controlled recordings in the classroom were conducted during the 2 year period from 1966 to 1968. This series of studies involved approximately 700 Ss including normals, mental retardates, and the culturally disadvantaged. The Ss were drawn from grades 3 through 6, and from classrooms for adolescent mental retardates. The experimental instructional materials were biographical passages of 2,000 to 4,000 words in length. Three of these passages comprised the set of materials known as the Standardized Listening Passages (Clark & Woodcock, 1967). The other studies utilized a set of 20 passages, each approximately 4,000 words in length, comprising the Negro Heritage Series (Clark, 1968). The several studies utilizing these materials were designed to approximate school learning situations as closely as possible while retaining the necessary research controls.

The results of these studies have provided information concerning several aspects of utilizing rate-controlled recordings in the classroom. The following questions were among those of concern:
1. Which type of media provides the most effective learning situation--listening plus viewing correlated slides, listening alone, or reading?
2. What is the effect of rate of presentation upon comprehension test scores and upon learning efficiency scores?
3. What is the relationship of intelligence to learning information through rate-controlled presentations?
4. How well do pupils retain the information learned via rate-controlled recordings as a function of time?
5. What are the effects of review upon performance?
6. What are the effects of practice with rate-controlled recordings upon performance?

---

*Dr. Richard W. Woodcock is Director of Research, American Guidance Service, Inc., Circle Pines, Minnesota 55014.

For the purposes of this paper, the details of research design and statistical analysis for the individual studies will not be presented. If this information is desired, it can be obtained from the reports listed in the references. (Clark, 1968; Clark & Woodcock, 1967; Woodcock & Clark, 1968a, 1968b, 1968c, 1968d.)

Studies Using the Standardized Listening Passages

Figure 9.1 illustrates a typical design of several studies which utilized the Standardized Listening Passages (Woodcock & Clark, 1968a). These studies typically examined performance across several words per minute (wpm) rates; compared immediate retention with retention after one week; and, generally, considered some aspect of intelligence. The Standardized Listening Passages procedure requires Ss to listen to a series of three passages. The first two passages and related tests are used for training and adaptation purposes; the third passage and its two alternate-form tests are used to provide experimental data.

Which Type of Media Provides the Most Effective Learning Situation?

The results of the studies reported in this paper were obtained through a presentation procedure in which Ss viewed a set of correlated slides while listening to a rate-controlled recording. During earlier studies, a marked improvement in performance was observed when Ss viewed correlated slides while listening to the presentation. One such study involved adolescent mental retardates who, in certain cases, read the passages using a controlled reading device; in other cases, listened to the passages; and, under the third condition, listened to the passages while viewing the correlated slides (Woodcock & Clark, 1968d). The results of this study demonstrated the superiority of the listening plus slide presentation over listening alone which, in turn, was more efficient as a learning medium than reading the passages.

What Is the Effect of Presentation Rate upon Comprehension Test Scores?

Figure 9.2 illustrates performance across several rates for a group of fifth-grade children with average intelligence. These data were derived from multiple-choice tests over the passage contents and were obtained immediately after the Ss listened to the passage and viewed the correlated slides. Note that higher performance is at the lower wpm rates, while performance at high wpm rates is relatively low. Performance at 378 and 428 wpm rates is at a level approximating that of Ss who took the test only without listening to the passages and viewing the slides.

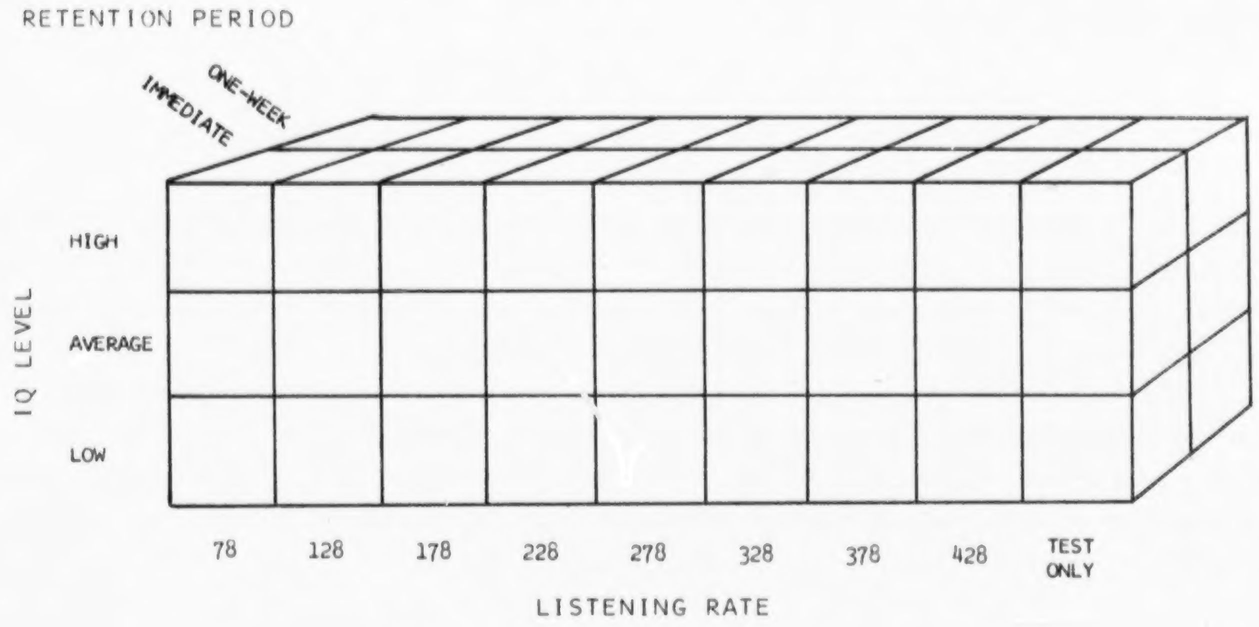

RETENTION PERIOD
ONE-WEEK
IMMEDIATE
IQ LEVEL
HIGH
AVERAGE
LOW
78
128
178
228
278
328
378
428
TEST ONLY
LISTENING RATE

Figure 9.1. Typical research design utilizing the Standardized Listening Passages.

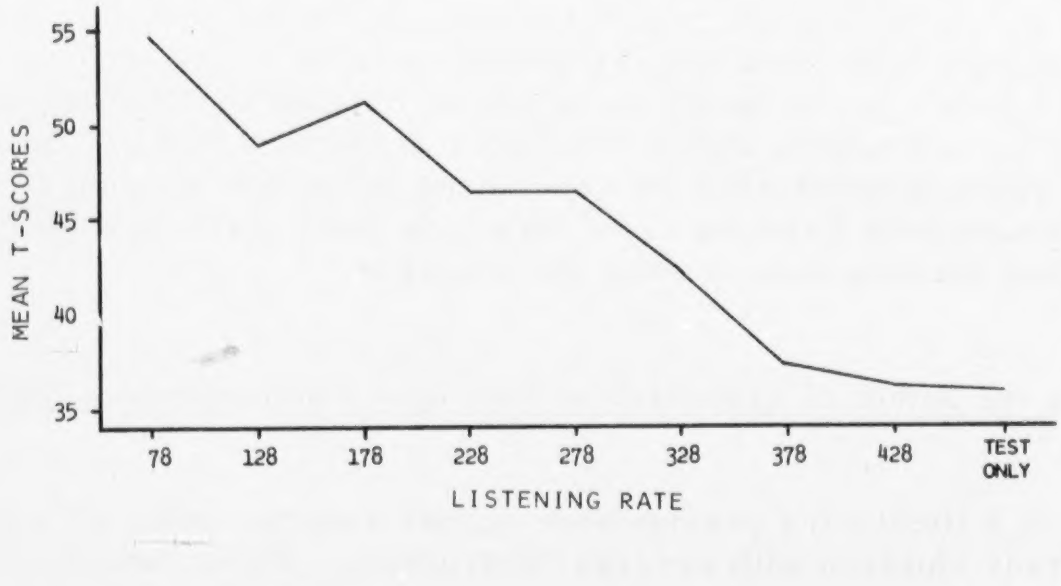


Figure 9.2. Performance across rate for fifth-grade pupils with average intelligence.

How Efficient Is Learning as a Function of Time Spent in Listening?

Figure 9.2 illustrated the effect of rate upon test scores. Figure 9.3 represents a second, and quite important, way to look at the effect of listening rate upon performance. The performance curve in Figure 9.3 portrays "learning efficiency" or the relative amount of learning per unit of listening time. Figure 9.3 illustrates that even though higher scores were made at lower wpm rates, more efficient learning takes place at higher wpm rates.

How Well Do Pupils Retain the Information Learned via Rate-Controlled Recordings?

A comparison of immediate retention scores with retention scores after 1 week are shown in Figure 9.4. Note that the 1 week retention curve is similar to the immediate retention curve except it is somewhat lower. The results portrayed in Figure 9.4 illustrate that information learned through the medium of controlled-rate recordings is retained and forgotten in much the same way as has been observed for learning obtained through other modes.

What is the Relationship of Intelligence to Learning via Rate-Controlled Recording?

Figure 9.5 illustrates performance curves for three levels of intelligence within the fifth grade. The curve representing the highest level of performance across listening rates was obtained by fifth-grade Ss with high mental ages. The middle curve is the same data shown in Figure 9.2 for average fifth-grade Ss. The lowest curve portrays the performance of fifth-grade Ss with low mental ages. Note the similarity of the three curves, except for their general level of performance.

Another aspect of the relationship of intelligence to learning performance is shown in Figure 9.6. In this case, the Ss comprising the three intelligence levels have the same mental age but differ in IQ and chronological age. The performance curve for the average fifth-grade Ss is again the same data as was shown in Figure 9.2. The low ability Ss are "below-average IQ" students in grade 6. The bright Ss are "above-average IQ" Ss in grade 3. The results of this study indicate that when Ss have the same mental age, even though they have different chronological ages and IQs, their performance will be similar.

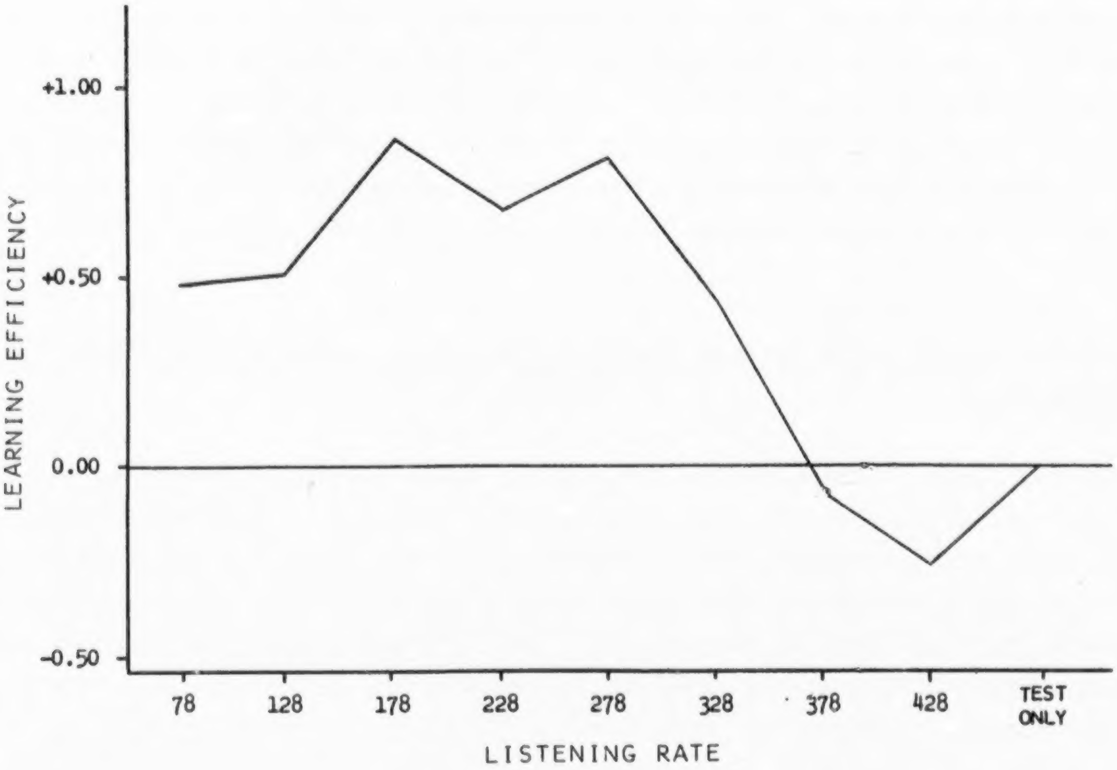


Figure 9.3. Learning efficiency across rate for fifth-grade pupils with average intelligence.

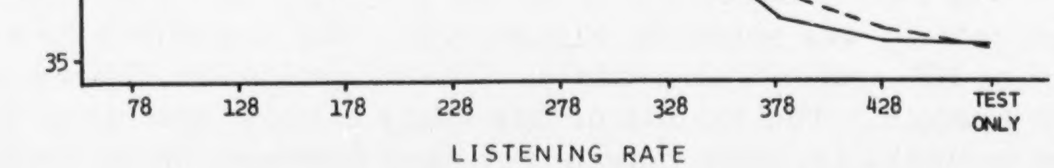


Figure 9.4. Comparison of immediate with one-week retention scores for average fifth-grade pupils.

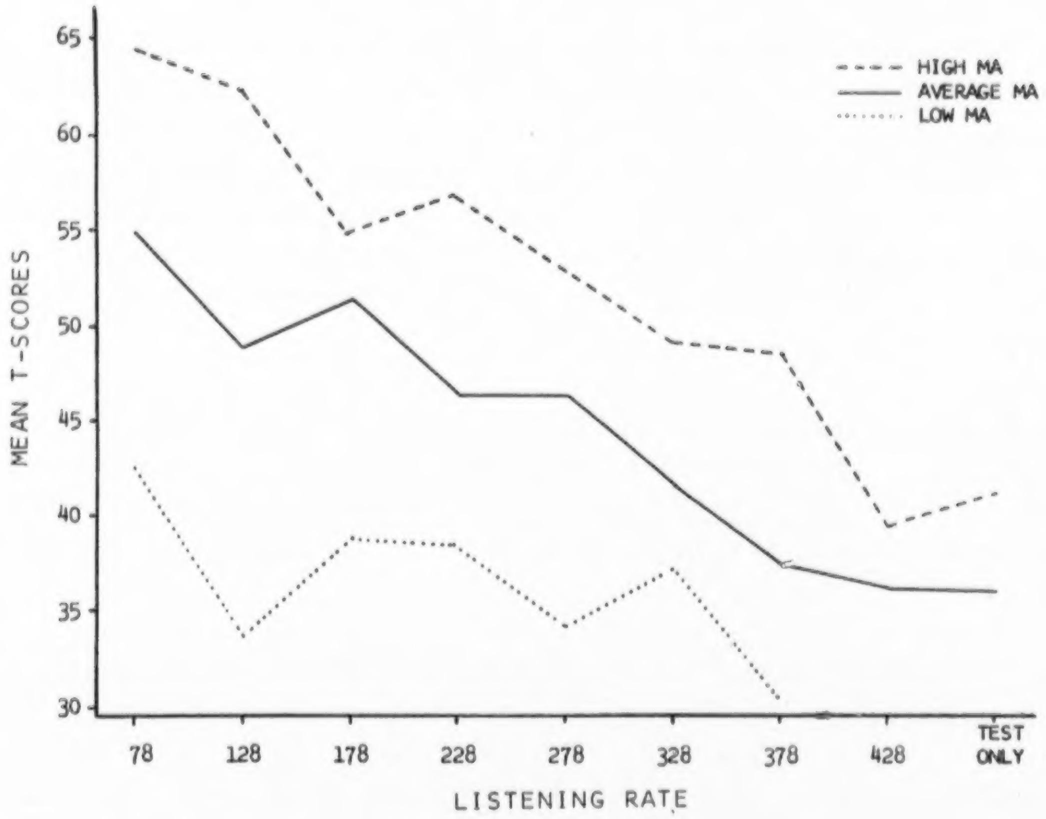


Figure 9.5. Comparison of immediate retention scores for fifth-grade pupils from three levels of intelligence.

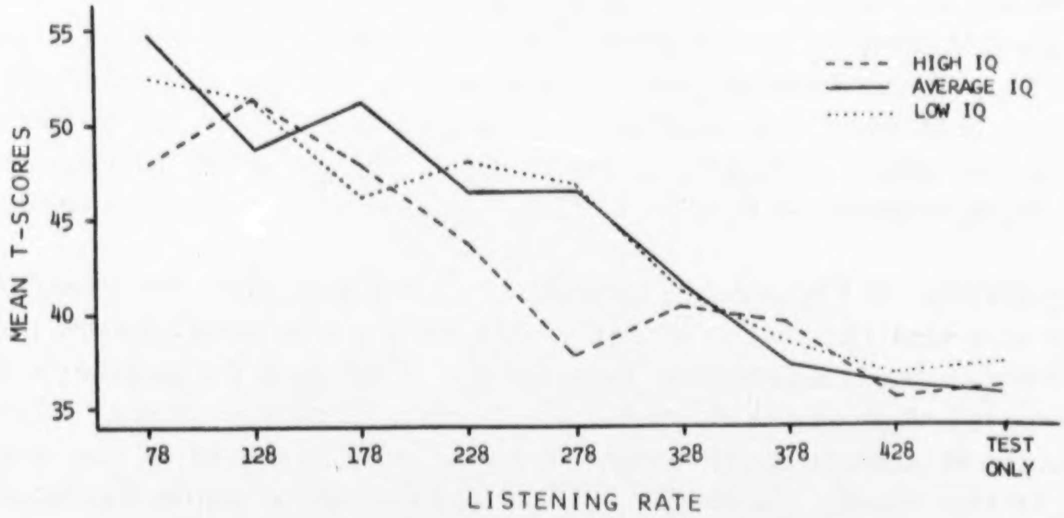


Figure 9.6. Comparison of immediate retention scores for pupils with the same mental age but from three levels of IQ.

What Are the Effects of Review upon Performance?

Figure 9.7 illustrates the design of a study which compared: (1) the single presentation of a passage with, (2) the double presentation of the passage with, (3) a variation of the double presentation strategy in which 1 week elapsed between the first and second exposures to the passage (Woodcock & Clark, 1968b). A question of special interest is whether Ss who listen to the material twice, at double the rate, will demonstrate more learning than those Ss who spend the same total listening time, but only listen to the passage once.

Figure 9.8 summarizes the results of the review study in respect to test performance. The overall differences among the three presentation strategies were not significant at the .05 level, though it appears that at very fast rates there may be a slight advantage in listening to the material twice.

Figure 9.9 portrays the results of the review study in respect to learning efficiency curves. Note that the results are similar for all three approaches. There is a trend in these curves suggesting that the single presentation may be slightly more efficient at slower rates and that double presentations may be more efficient at higher rates.

Studies Using the Negro Heritage Series

Figure 9.10 presents the design of the major study reported in this paper involving the application of rate-controlled recordings in the classroom (Woodcock & Clark, 1968c). The 20 Negro Heritage passages each required about 20 minutes of listening time at their original recorded rate. The Ss listened to one of these passages each day, 4 days a week, for 5 weeks. After listening to each passage, the Ss were administered a 12-item test over the contents of the passage. One week after completing the twentieth passage in the series, the Ss were administered an 80-item comprehensive test over the contents of all 20 passages.

Examination of Figures 9.11 and 9.12 indicate that the results of this study are similar to the results obtained in the short-term studies using the Standardized Listening Passages. Figure 9.11 portrays the comprehensive test score data for the Negro Heritage study. Note that again, the highest scores were obtained by Ss who listened at the slowest rate used in the study, while the lowest scores were made by those Ss who listened at the fastest rate. The only Ss who did poorer were those who took the test without having had the benefit of listening to the passages. Figure 9.11 also compares the performance of Ss who listened to each passage twice with those Ss who listened only once. As seen before, in Figure 9.7, there seems to be little advantage from listening to the passages twice.

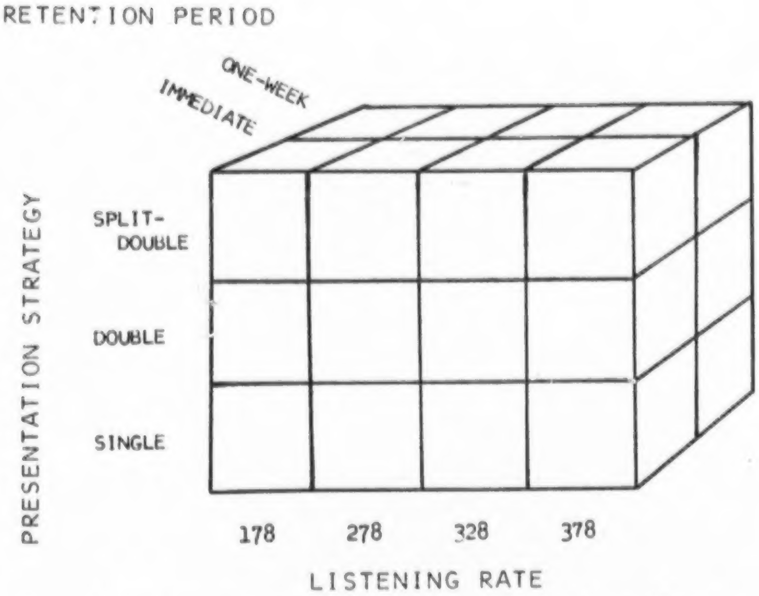

RETENTION PERIOD
ONE-WEEK
IMMEDIATE
PRESENTATION STRATEGY
SPLIT-DOUBLE
DOUBLE
SINGLE
178
278
328
378
LISTENING RATE

Figure 9.7. Design of the study comparing three presentation strategies.

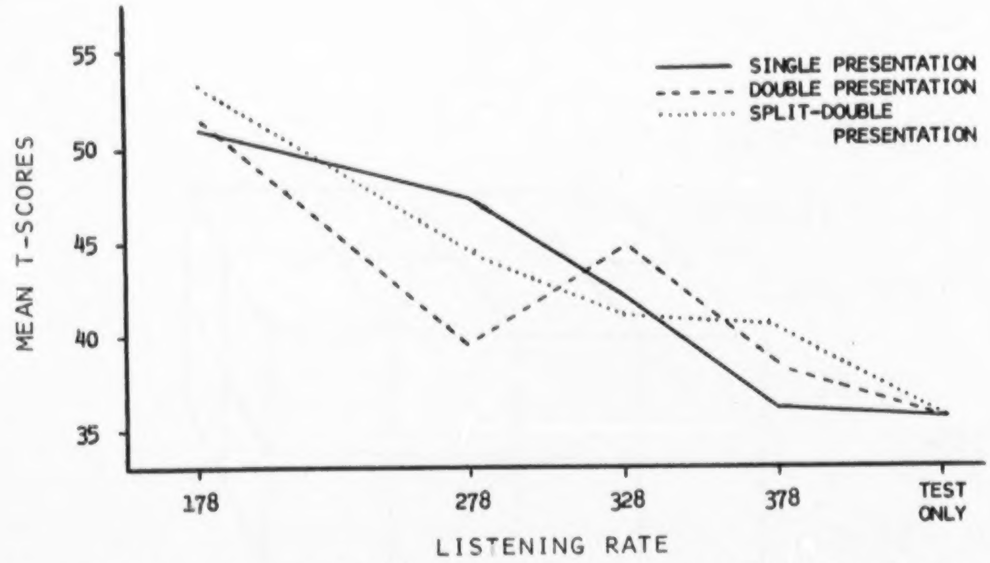


Figure 9.8. Immediate retention scores for the three presentation strategies.

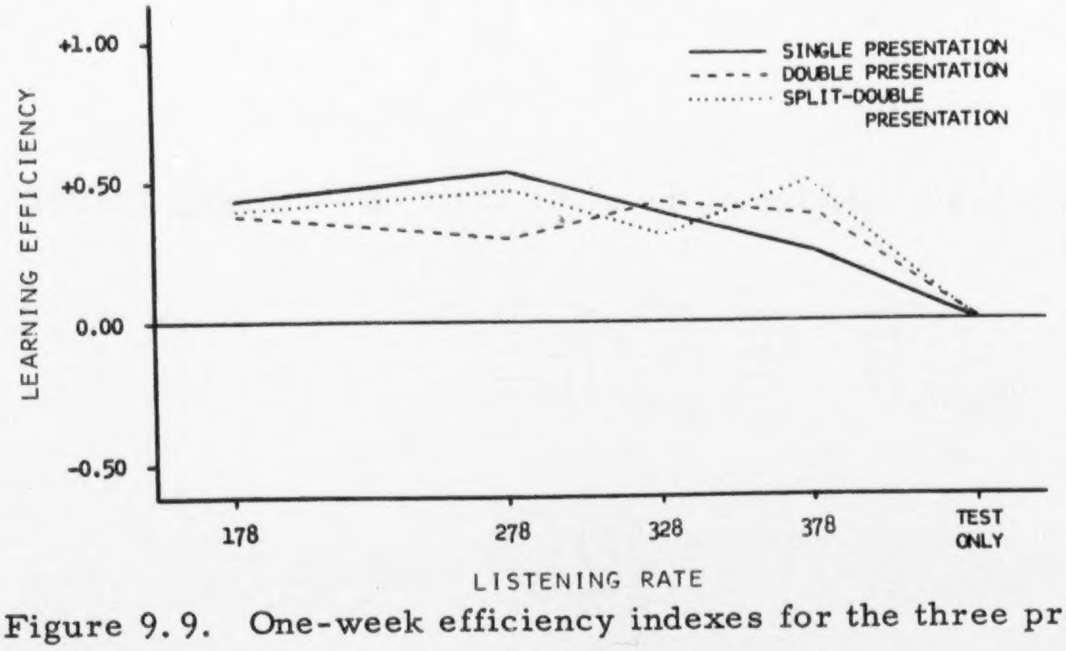


Figure 9.9. One-week efficiency indexes for the three presentation strategies.

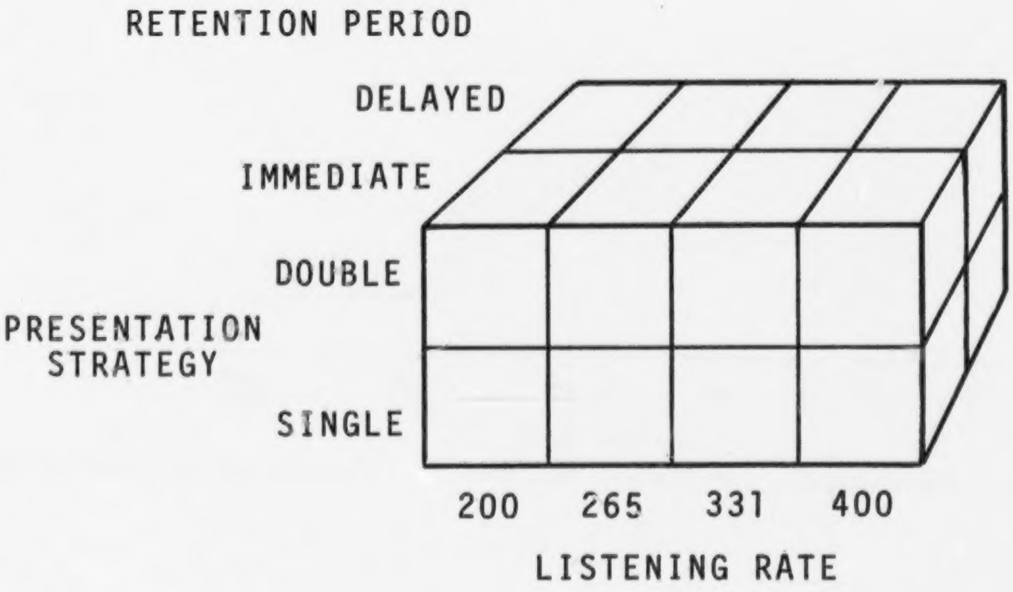

RETENTION PERIOD
DELAYED
IMMEDIATE
DOUBLE
PRESENTATION STRATEGY
SINGLE
200
265
331
400
LISTENING RATE

Figure 9.10. Design of the study utilizing the Negro Heritage passages.

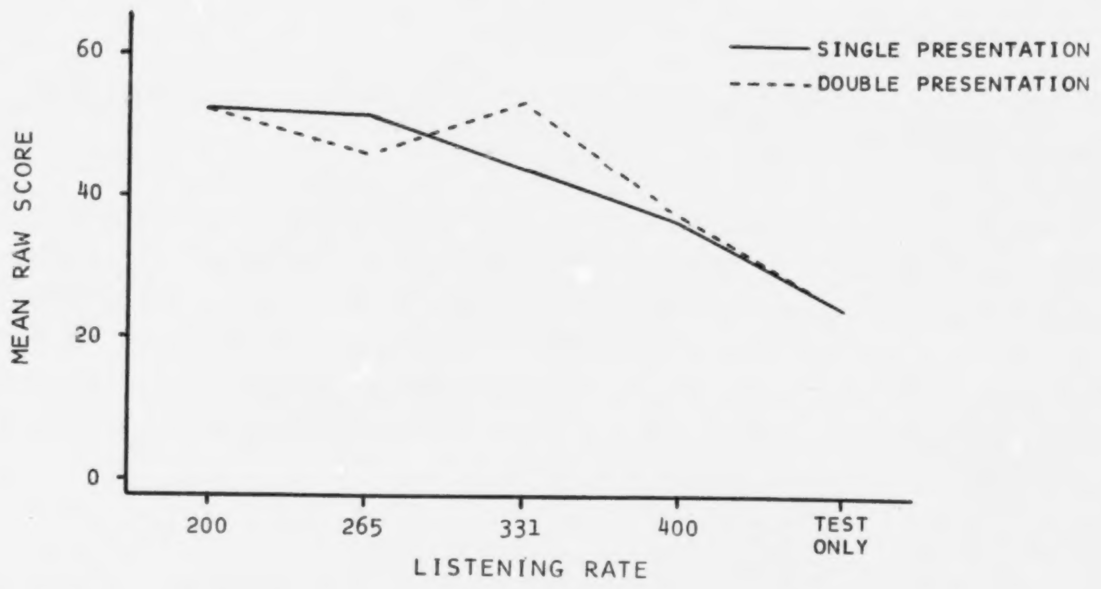


Figure 9.11. Delayed retention scores for the Negro Heritage study.

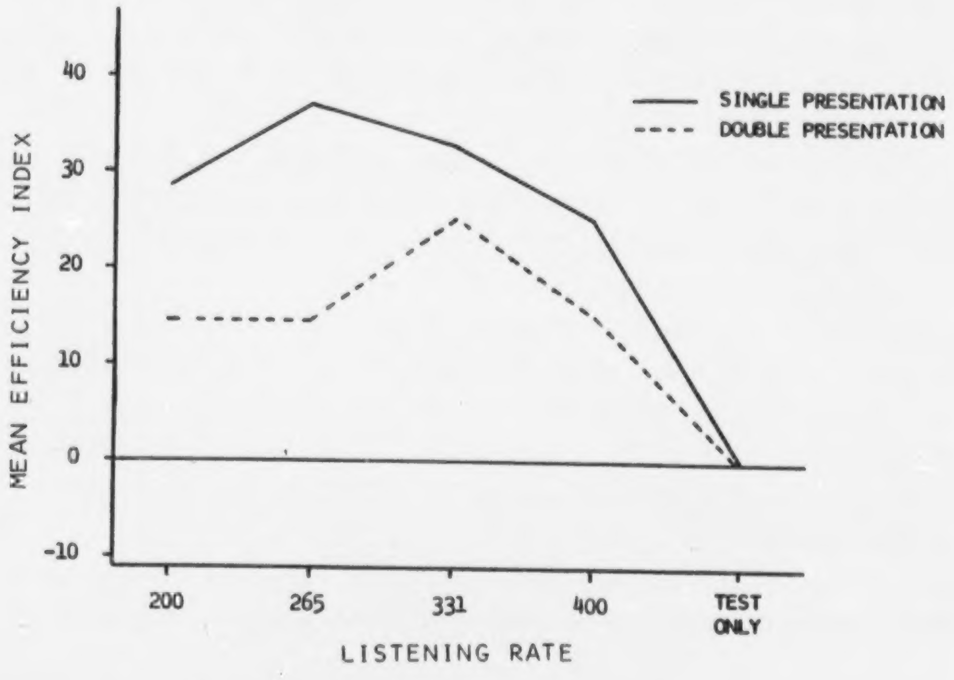


Figure 9.12. Learning efficiency scores for the Negro Heritage study.

Figure 9.12 shows the data of Figure 9.11 transformed into learning efficiency scores. The loss in learning efficiency by listening to the same passage twice is clearly demonstrated. At no point did the efficiency of learning for those Ss who listened twice approach that of Ss who listened only once.

What Effect Does Practice with Rate-Controlled Recordings Have upon Performance?

Figure 9.13 portrays the week-by-week results obtained in the Negro Heritage study. It appears that very little change in performance took place after the initial exposure to rate-controlled recordings. A limiting factor in this study was the lack of feedback to Ss regarding their week-to-week performance; however, in respect to sheer exposure to the medium there was no marked improvement in performance with practice.

Conclusions and Summary

Six major conclusions regarding the application of rate-controlled recordings in the classroom are implied by the results of the studies presented in this paper:

1. Listening plus viewing slides is a more effective and efficient medium for learning than is listening alone, which, in turn, is more effective than reading (at least for Ss who are not yet good readers).
2. A pupil will achieve the highest score on a test over a passage at expanded rates of 75 to 125 wpm.
3. A pupil's most efficient learning will take place at compressed rates of approximately 250 to 300 wpm. (It is of interest to note that the normal speaking rate of 150 to 175 wpm provides neither the most effective rate nor the most efficient rate.)
4. In respect to the relationship of performance to intelligence, mental age is a very significant S variable. IQ, when mental age is held constant, does not seem to be an important variable.
5. In respect to the value of review, a single pass through the material is a more efficient use of available learning time than repeated exposures to the same material.
6. After the initial two or three exposures to rate-controlled recordings, continued practice produces little improvement in performance.

In summary, if I were assigned the task today of setting up an extended program of school instruction using controlled-rate recordings, I would do the following: prepare lessons to be presented audio-visually, not audio alone. The visual presentation would change every one-half to 1 minute. If the instructional goal is for pupils to achieve the highest

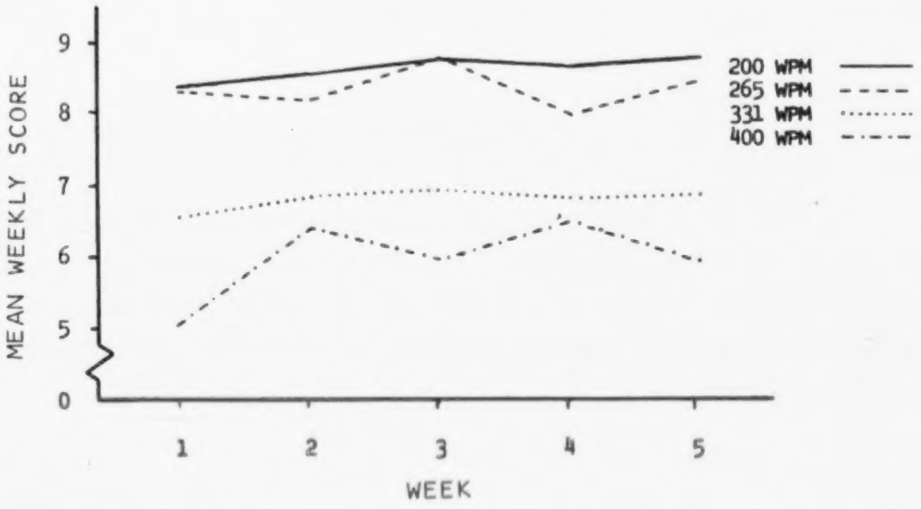


Figure 9.13. Effect of practice upon weekly scores (single presentation only).

score possible after listening to the oral presentation, they would listen once at a rate of 75 to 100 wpm. If the instructional goal is to achieve the most learning in a limited time, the pupils would listen to the passage once at a rate of 250 to 300 wpm. Following each presentation, the pupils would be administered a short test or other evaluative device over the contents and implications of the material to which they have just been exposed.

These studies have provided further evidence that rate-controlled recordings can be an effective and efficient learning medium for children. The self-contained nature of such materials and of "listening-viewing centers" can provide teachers with an easily handled and effective instructional situation.

ACKNOWLEDGMENTS

The research reported herein was supported by the Institute on Mental Retardation and Intellectual Development (George Peabody College) under a grant from the National Institute of Child Health and Development, and by the Nashville Education Improvement Project under a grant from the Ford Foundation.

CHAPTER X

PROCESSING TIME AS A VARIABLE IN THE COMPREHENSION OF TIME-COMPRESSED SPEECH

Ruth Ann Overmann*

Time-compressed speech is speech reproduced in less than the original production time. The compression process can be accomplished in a number of different ways, but the method used in the work to be described is one that permits an increase in word rate without altering other characteristics of the message such as timbre or pitch. The determination of the optimal rate of speech at which man can still adequately encode, store, and retrieve the information in a spoken message is a problem of great interest for those engaged in research on the perception of time-compressed speech and it is to this problem that the following paper is directed.

Originally, interest in time-compressed speech grew out of a need to increase the amount of information that could be communicated aurally in the time available for such communication. Blind students, for example, would be significantly assisted if they could read at a rate that compares favorably with the silent visual reading rate of 251 words per minute (wpm), the median reading rate for high school seniors (Harris, 1947). The reading rate of the individual who reads by listening is ordinarily determined, not by his own reading requirements, but by the rate at which his oral reader speaks. Foulke (1969, Ch. 11) found a mean oral reading rate of 177 wpm for a representative group of professional oral readers employed in the Talking Book program. Typical braille reading rates are even slower. Ethington (1956) found an average braille reading rate of 90 wpm for high school students, and 120 wpm for experienced adult braille readers. Thus, under the best of conditions, the blind reader receives two-thirds the amount of information that the sighted reader receives in the same period of time.

---

*This experiment was performed by Ruth Ann Overmann while a graduate research assistant in the Perceptual Alternatives Laboratory at the University of Louisville, Louisville, Kentucky, and was reported in her master's thesis.

A method for the time compression of recorded speech without the distortion in vocal timbre and pitch associated with the accelerated playback of a tape or record was suggested by the results of an experiment reported by Miller and Licklider (1950). They demonstrated the redundancy of normal speech by testing the intelligibility of monosyllabic words, the reproduction of which was interrupted periodically by means of an electronic switch. They found that when the samples eliminated from the reproduction by interruption of the signal did not exceed 50 milliseconds (msec.) in length, word intelligibility did not fall below 90% until 50% of the signal had been eliminated. In this study, intelligibility was operationally defined as the ability to repeat words accurately, and this definition is followed in the work reported here.

The next step in the development of compressed speech was to close the gaps between the remaining parts of the interrupted words. Garvey (1953b) accomplished this operation by periodically cutting out segments of a recorded tape, and by splicing the free ends together again. The speech that resulted from the reproduction of this tape was accelerated without distortion in vocal pitch. Obviously, this technique of speech compression was far too time-consuming and cumbersome to be of any practical use, but it did demonstrate the soundness of the approach.

In 1954, Fairbanks, Everitt, and Jaeger described an electromechanical device for the time compression of speech. This device makes use of a sampling wheel, with four tape playback heads equally spaced around its curved surface. As the tape to be reproduced passes over the curved surface of the sampling wheel, the wheel rotates in the direction of tape motion and, as each playback head is periodically brought into contact with the moving tape, it reproduces a sample of the signal recorded on the tape. As a given head loses contact with the tape and, hence, ceases to reproduce the signal recorded on it, the next head in line makes contact and commences reproducing. However, a segment of tape, equal in length to the distance along the curved surface of the sampling wheel separating these heads, is skipped, and is therefore eliminated from the reproduction. The output, then, consists of samples of the original recording, abutted in time. The amount of compression depends upon the frequency with which segments of the original recording are discarded. Speech intelligibility will be preserved if the discarded samples are of short enough duration so that every speech sound is sampled (Foulke, 1969, Ch. 2).

Using this method, it is possible to accelerate the word rate of connected discourse by any desired amount. The basic limitation in amount of compression is the listener's ability to comprehend what he hears.

The effect of compression on the intelligibility of single words and on the comprehension of connected discourse has been studied by several investigators. Garvey (1953b), using his manual sampling method, found a 10% loss of intelligibility for words compressed to 60% of original duration. Compressing fluent speech by this would result in a word rate of 292 wpm, assuming an initial word rate of 175 wpm. Kurtzrock (1957), using the electromechanical sampling method, found 50% intelligibility for words reproduced in only 15% of original production time. Fairbanks and Kodman (1957) found 50% intelligibility when only 13% of the original word was present. Connected discourse, subjected to a compression of this magnitude, would be reproduced at a rate of 1,000 wpm, or more. The low intelligibility of the individual words in messages presented at these rates should be counteracted to a considerable degree by the redundancy in connected discourse, and theoretically, it should therefore be possible to demonstrate significant comprehension of messages presented at rates in the range of 1,000 wpm. Yet, as the following research will clearly indicate, this is not the case.

Comprehension, as measured by performance on a test of the facts and implications of a listening selection, has also been shown to decrease as a function of an increase in word rate. Both Nelson (1948) and Harwood (1955) demonstrated an insignificant loss in comprehension for word rates in the range of 125 to 225 wpm. For listening selections of 141, 201, and 282 wpm, Fairbanks, Guttman, and Miron (1957c) found little difference in comprehension scores. However, comprehension declined from 58% of test items correctly answered at 282 wpm to 26% at 470 wpm. Since the tests used were of the multiple-choice type, a mean score of 26% would not be significantly different from chance performance.

The nature of the material presented in a listening selection could be one of the factors determining the rate at which comprehension is lost as word rate is increased. Using both literary and technical listening material, Foulke, Amster, Nolan, and Bixler (1962) demonstrated that comprehension was only slightly affected by word rates of up to 275 wpm. In the range from 275 wpm to 375 wpm, there was a sharp decline in comprehension as the word rate was increased for both types of listening material. Also, the comprehension scores of the blind Ss at the 275 wpm rate were essentially the same as the scores of the sighted Ss who had silently read the material and answered the test items.

In a single experiment, Foulke and Sticht (1967a) measured the influence of compression on both word intelligibility and listening comprehension. They found that both intelligibility and comprehension decreased as the amount of compression was increased but that comprehension declined more rapidly than intelligibility. The intelligibility of single words was always superior to the comprehension of connected discourse, and declined gradually as the compression was increased from the value required for a word rate of 225 wpm to the value required for a word rate of 425 wpm. Comprehension scores, on the other hand, declined moderately from 225 to 325 wpm, but very rapidly thereafter.

The conclusion suggested by these and other studies is that listening comprehension begins its rapid decline at a compression rate that leaves word intelligibility relatively intact. It is, of course, to be expected that comprehension scores would be lower than intelligibility scores. The listener's task, when he is required to demonstrate comprehension, is more difficult than the task required when intelligibility is measured. However, if the decline in listening comprehension is due solely to the loss of word intelligibility, then an improvement in the intelligibility of words presented at a given level of compression should be followed by the improved comprehension of messages constructed from those words. In an experiment by Foulke (1969, Ch. 12), listeners were taught a vocabulary of words compressed by an amount sufficient to produce a rate of approximately 500 wpm for connected discourse. They received practice in the identification of these words until they had achieved near perfect performance. Then, they were required to reproduce sentences composed of these words and presented at approximately 500 wpm. Errors in reproduction were numerous and did not yield to practice in spite of the steps that had been taken to insure the intelligibility of component words.

Another source of evidence for the rejection of poor word intelligibility as the only factor accounting for the decline in listening comprehension at high levels of compression is provided in an experiment by Foulke (1969, Ch. 12). A professional reader read the same listening selection at three different word rates: 149, 164.6, and 195.7 wpm. These three renditions were then compressed to the same final rate of 275 wpm. In order to achieve this, it was necessary to compress the three renditions to 71%, 60%, and 50% respectively of their original production times. In spite of the fact that the intelligibility of the words in the three renditions differed as a consequence of the different amounts of compression to which they had been subjected, the distributions of comprehension test scores for the three groups of Ss who heard these selections were not significantly different.

In any case, the finding that intelligibility is lost at a different rate than comprehension as the amount of compression is increased remains to be explained. In one attempt to explain this fact, Foulke and Sticht (1967a) have adopted the computer storage model of cognitive processing. This model describes in terms of input, matching, storing, and retrieval operations, the task of reading a verbal selection and remembering significant aspects of it. The conscious sensory acts involved in reading or listening are equivalent to the input operations of feeding data into a computer. The next operation, storing this material, involves two processes: first, the matching of the words in the message with an already stored vocabulary of words and phrases; and second, the as yet unspecified mechanisms necessary for the synthesizing and coding of the material for memory. Memory is of two types, long-term and short-term. Presumably, only the whole intact message is involved in short-term memory operations and it is in this brief time that the matching and coding of the message takes place for the long-term storage. Retrieval is the name given to the actual remembering process, in which the coded information is decoded.

Thus, as Foulke and Sticht (1967a) point out, word intelligibility could be explained in the operation sequence of: input, matching short-term memory storage, and retrieval. However, the task is not quite so simple in comprehension. There must be a continuous process of input, matching, buffer storage, those encoding processes required for the transduction of input material to a form suitable for long-term storage, and a final decoding step required to retrieve information from the long-term memory storage bank.

The significant difference between the two tasks is the added number of processing operations. In normal speech of approximately 150 to 200 wpm, there is apparently more than enough time to perform all of these processing operations on all of the incoming information to make what is heard understandable. The individual can still retrieve the message adequately as evidenced by the relatively high scores on tests of comprehension. However, as the word rate of listening material is increased, the time available for these processing operations is decreased. A rate is ultimately reached at which there is no surplus time in which to perform the needed operations. As word rate is increased beyond this point, the rapid decline in comprehension begins. As seen in the previous studies, the individual is still able to handle rates of up to 275 wpm; but after this point is reached, comprehension scores begin to decline.

To explain the information processing capacity of the organism, Miller (1953, 1956) has employed the concept of a communication channel which has a finite capacity. If more information is presented to the individual than he can handle, then some of this information will not be processed and cannot subsequently be stored for later retrieval. Thus, in higher word rates, the channel capacity could be exceeded and information would be lost. Compressed speech complicates this process all the more. In the language of the computer storage model, there are fewer cues present in the compressed word to aid in the recognition process and consequently more items in the individual's store of vocabulary must be rejected before the correct match is found. Essentially, the process of compression reduces the redundancy of the individual words in the message that is to be comprehended. Thus, as available processing time decreases with increasing word rate, the growth in word uncertainty increases the demand for processing time. After channel capacity is reached at approximately 275 wpm, comprehension should begin to decline, and the slope of this line should become gradually steeper as the rate of compression is increased. Foulke (1968b) compared the results of several investigations of comprehension and the evidence offers tentative support for this hypothesis.

If the suggested explanation of the change in the rate at which comprehension is lost with increasing word rate is correct, one consequence should be that efforts toward training listeners to comprehend very fast speech will meet with only limited success. The nature of the neural mechanisms involved in the processing of stimulation may set a fairly inflexible upper limit on the rate at which units of information can be obtained from stimulus displays, and efforts to exceed this limit may only interfere with the listener's perceptual machinery. The results of training efforts so far attempted seem to confirm this expectation. Foulke (1964a) evaluated the efficiency of four different training methods in improving the comprehension of compressed speech. One group received prolonged practice in listening at a constant, high word rate. A second group received practice in listening to material presented at a word rate that was initially slow, but which was gradually increased to a very fast word rate over the course of training. For two additional groups, word rates were managed in the same way, but listening selections were interrupted frequently to question Ss about material just heard. None of the four training methods yielded any significant improvement in the comprehension of compressed speech.

Other attempts at improving the comprehension of compressed speech by training have been only moderately successful (Orr, Friedman, & Williams, 1965). There does appear to be an initial "warm-up" effect, or adjustment to the task of listening to compressed speech, as shown in the results of an experiment by Voor and Miller (1965). They presented five different listening selections to a group of Ss at a rate of 380 wpm. Each selection was followed by a multiple-choice test of comprehension. There was an improvement in mean comprehension scores from the first to the third selection, but from the third to the fifth selection, there was no further improvement.

If the influence of time compression on the comprehension of connected discourse is consonant with the model suggested in this paper, then the reinsertion, at the syntactic boundaries in a listening selection, of the processing time removed by compression should restore the comprehensibility of that selection. There already is some support for this hypothesis in the literature. Aaronson and Markowitz (1967) studied the effects of compression on the recall of sequences of spoken digits. In one condition, Ss listened to the straight reproduction of the recorded oral reading of sequences of numbers. In the other condition, these recorded sequences of numbers were compressed to 67% of original duration, and enough unfilled time was inserted between each of the spoken numbers to restore the time required for the reproduction of the sequence to the original production time. The results of this experiment showed that immediate recall accuracy was significantly higher for compressed than for noncompressed sequences, thus suggesting that the additional information in the normal sequences was not necessary for encoding and retrieval, and that processing time was the significant variable.

The question of the relative contribution of degradation in word intelligibility and deprivation of processing time to the decline in comprehension of highly compressed speech is an important issue for both theoretical and practical reasons. Theoretically, its resolution would contribute to an increased understanding of the perception of verbal inputs. Practically, if the problem is primarily one of poor word intelligibility rather than insufficient channel capacity, attention can be directed toward the development of training methods to help the listener discriminate sounds that are unfamiliar in their compressed form, and to the development of improved equipment for the compression of speech. If insufficient channel capacity is the principal difficulty, one might consider the investigation of various strategies that could be employed in listening to compressed speech.

The following research was an attempt to answer the question of why comprehension declines at high word rates of compression by suggesting that the loss is due to a limited channel capacity for processing information and not to a loss of word intelligibility. An experiment was performed using tapes that were time-compressed at different word rates and tapes in which the individual phrases and sentences were compressed with time added to restore the tape to the original production time. This unfilled time was to replace the processing time lost in the compression process. If the adding of processing time is the answer to the loss of comprehension, the comprehension scores should parallel the intelligibility scores at all word rates. In the following experiment, if this hypothesis is true, the selections to which processing time has been added should result in: (1) significantly higher comprehension scores than the time-compressed selections that have not had any time added to them, and (2) a significant interaction effect that would indicate the different rates of decline in comprehension as the compressed word rate increases for the two groups.

Method

Subjects

One hundred and forty introductory psychology students at the University of Louisville served as Ss in this experiment. None had prior experience with compressed speech and all were free from any obvious hearing defects.

Experimental Materials

Listening comprehension was measured by the Nelson-Denny Test of Reading Comprehension, Form B. Form B consists of eight short listening selections with several questions covering the material in each selection. Selection 1 contains 600 words with eight questions, and each of the remaining selections contains approximately 200 words, with four questions per selection. A score of 36 is the highest possible score.

The selections were recorded by a professional reader at the American Printing House for the Blind. Two copies were made of this tape. One tape was compressed using an electromechanical speech compressor of the Fairbanks type (Fairbanks et al., 1954) built at the University of Louisville. The tape was reproduced by the speech compressor at those compressions needed to achieve word rates of 250, 325, and 400 wpm. Playback time was measured for each selection and the differences between the compressed form and the original playback time was figured for each of the three compressed word rates. The number of pauses between phrases and sentences was then determined for each reading selection. The time difference between each of the compressed word rates and the normal word rate was distributed between the phrases and sentences with the pauses between the sentences receiving twice as much time as the pauses between the phrases. The time was added to the tape by splicing in the proper amount of leader tape between the phrases and sentences and the resulting tape was copied for use in the experiment. This tape contained reading selections with phrases and sentences compressed to the equivalent of 250, 325, and 400 wpm in connected discourse with enough unfilled time added to approximate the time of the original master tape playback. Table 10.1 is a compilation of playback times for the original compressed tapes and time-added tapes for each selection and for each compression rate.

The output of the speech compressor was recorded on magnetic tape via a Crown tape recorder, model 800. During the experiment, these tapes were reproduced on a Wollensak tape recorder, model T-1500. The output of this recorder was distributed to the Ss through cushioned headsets. Each headset's volume could be adjusted by the S for comfortable listening.

Procedure

The 140 Ss were divided into six experimental groups of 20 members each and one control group of 20. Each of the experimental groups received one of the compressed tapes--three heard the normally compressed tapes, and three heard the time-added compressed tapes. The control

TABLE 10.1

ORIGINAL PRODUCTION TIMES AND COMPRESSED REPRODUCTION TIMES

Selection	Total Number of Words	Number of Sentences	Number of Phrases	Original Production Time (in seconds)	Reproduction Time for 250 wpm (in seconds)	Difference Between Original and 250 wpm (in seconds)	Reproduction Time for 325 wpm (in seconds)	Difference Between Original and 325 wpm (in seconds)	Reproduction Time for 400 wpm (in seconds)	Difference Between Original and 400 wpm (in seconds)
1	625	30	63	197	150	47	115.4	81.6	93.75	103.25
2	200	8	20	61	48	13	36.91	24.2	29.31	32.69
3	201	7	21	62	48	14	36.81	25.2	30.33	31.67
4	200	6	26	69	47.5	20.5	37.92	32.19	29.82	39.18
5	223	13	18	72	53.5	22	41.15	30.83	33.45	38.5
6	204	10	21	71	49	22	37.65	33.34	30.6	39.4
7	205	15	22	74	48	25	37.84	36.16	30.6	43.4
8	201	8	20	62	50	12	36.81	25.2	30.14	32.06

group heard the uncompressed tape. The Ss were assigned to experimental groups in the order in which they signed for service in the experiment. All Ss were told that they would be listening to compressed speech, a form of accelerated speech in which word rate is increased without changing the "voice quality"; that they would be asked questions on each of the selections they were about to hear; and that they might have some difficulty understanding what was said but to try and do the best they could.

At the end of each selection, they were given a test containing the questions covering that selection and were instructed to read the questions and mark their answers on a numbered IBM answer sheet. After the Ss had completed each test, they listened to the next selection on the tape.

Results

A score, the number of correctly answered test items, was determined for each S. The means and standard deviations of these scores were calculated for each experimental group, and for the control group (see Table 10.2). The effect of time compression on listening comprehension, for the "time compression" groups, and for the "time compression plus time restoration" groups, is shown in Figure 10.1. The x-axis of this figure is scaled in terms of amount of compression, and the y-axis in terms of mean test scores. A dotted line has been drawn across the graph, parallel to the x-axis, to indicate the level of chance performance. The mean test scores of all experimental groups are clearly above chance at all three compressions. Furthermore, mean test scores for the "time compression plus time restoration" groups are higher than the mean test scores for the "time compression" groups.

TABLE 10.2

MEANS AND STANDARD DEVIATIONS

Control Group		"Time Compression" Groups			"Time Compression Plus Time Restoration" Groups		
M	SD	WPM	M	SD	WPM	M	SD
25.95	2.95	250	24.00	4.32	250	26.15	4.12
		325	22.50	3.85	325	23.55	5.92
		400	18.55	5.29	400	21.30	6.90

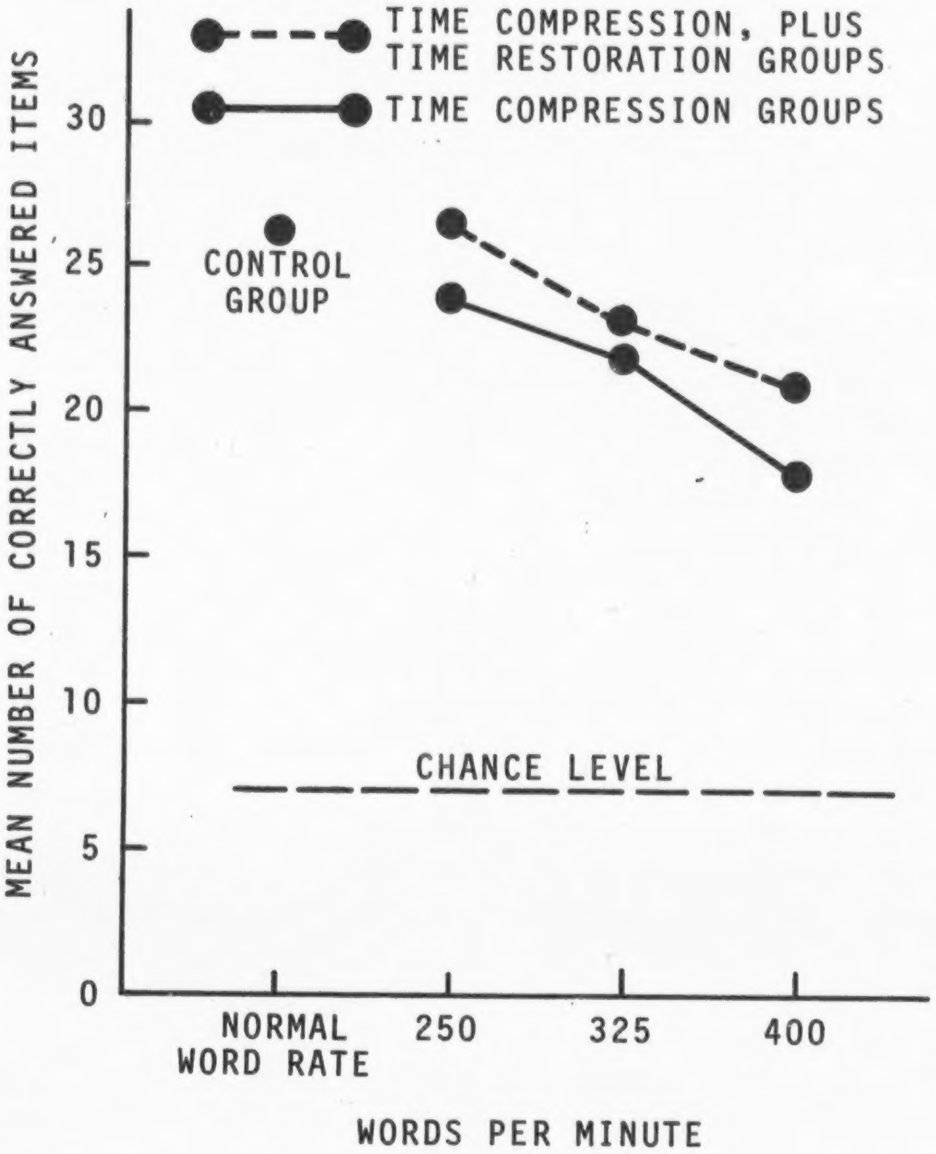


Figure 10.1. Comprehension test scores as a function of word rate.

The data upon which Figure 10.1 is based were examined by an analysis of variance, using a 3 x 2 factorial design with a single control group (Winer, 1962). The results of this analysis are presented in Table 10.3. As suggested in Figure 10.1, effects due to simple time compression, and to time compression plus time restoration, were significant ($p < .01$ and $p < .05$ respectively). The interaction between time compression and time restoration was not significant.

TABLE 10.3

ANALYSIS OF VARIANCE

Source	df	MS	F	p
Control vs all others	1	183.88	7.70	.01
A (Time Compression)	2	268.90	11.26	.01
B (Time Compression Plus Time Restoration)	1	118.01	4.94	.05
AB	2	7.43		
Within (pooled with control)	133	23.88		

Further analyses on these data were performed to determine the sources of significance. The Newman-Keuls Test for Ordered Pairs of Means (Winer, 1962, pp. 80-85) indicated that, for the "time compression" groups, performance at the compression required for a word rate of 250 wpm (assuming an uncompressed word rate of 175 wpm) was significantly different from performance at the compression required for a word rate of 325 wpm, and also significantly different from performance at the compression required for a word rate of 400 wpm, while performance at the compression required for a word rate of 325 wpm was significantly different from performance at the compression required for a word rate of 400 wpm ($p < .05$ in all cases). The same pattern of significance was found for the "time compression plus time restoration" groups.

In addition, at each of the three levels of compression represented in the experiment, the difference between the distributions of scores for the "time compression" and the "time compression plus time restoration" groups was examined for significance by the Newman-Keuls test. At both the compression required for a word rate of 250 wpm and the compression required for a word rate of 400 wpm, the scores of the "time compression" groups were significantly different from the scores of the "time compression plus time restoration" groups ($p < .05$ in both cases). However, at the compression required for a word rate of 325 wpm, the difference was not significant.

In a further analysis, the _t_ test was used to compare each experimental group with the control group. The results of this analysis are shown in Table 10.4. At the compression required for a word rate of 250 wpm, neither the "time compression" group, nor the "time compression plus time restoration" group, show a significant loss in listening comprehension. Beyond this point, the loss in comprehension is significant for the "time compression" group; whereas, for the "time compression plus time restoration" group, the loss in comprehension does not reach significance until the compression required for a word rate of 325 wpm is exceeded.

TABLE 10.4

_t_ RATIOS FOR COMPRESSED PLAYBACK TIME

WPM	"Time Compression" and Control Groups	"Time Compression Plus Time Restoration" and Control Groups
250	1.67	.18
325	2.61*	1.67
400	5.48*	2.77*

* $p < .01$

Discussion

As the results clearly indicate, both time compression and time restoration are significant factors in the comprehension of compressed speech. The demonstration of decreasing comprehension with increasing compression in time, constitutes a replication of earlier findings (Fairbanks et al., 1957c; Foulke, 1968a; Foulke et al., 1962; Foulke & Sticht, 1967a).

The finding that the restoration, at phrase and sentence boundaries, of the time lost by compression improved comprehension, supports the hypothesis that the loss in comprehension resulting from compression in time is a consequence of depriving the listener of the time he needs to perform the processing operations upon which comprehension depends. Although the effect of time restoration was, in all instances, to improve comprehension, the improvement was, in one instance, not significant. At the compression required for a word rate of approximately 325 wpm without time restoration, it would have been necessary for the mean comprehension test score of the "time compression plus time restoration" group to exceed the mean comprehension test score for the "time compression" group by 1.17. The observed difference was only 1.06. However, since the difference produced by time restoration was in the expected direction in all three cases, and significant in two of the three cases, the failure to realize significance in the remaining case does not seriously weaken the time restoration hypothesis.

As shown in Figure 10.1, the comprehension of compressed speech was not significantly poorer than the comprehension of uncompressed speech until a word rate of approximately 250 wpm was exceeded. This finding is in general agreement with the findings of other studies of comprehension as a function of word rate, cited earlier. If the restoration of the time loss by compression had been completely effective, there would have been no loss in comprehension with increasing compression for the "time compression plus time restoration" groups. However, as shown in Figure 10.1, although the comprehension manifested by the "time compression plus time restoration" groups did not decline significantly until the compression required for a word rate of approximately 325 wpm without time restoration was exceeded, the loss in comprehension was significant at the highest compression represented in the experiment. One explanation may be that, at this compression, comprehension was adversely affected, in part, by a degradation of signal quality. Of course, the restoration of the time lost by compression would not compensate for such degradation. Another possibility is that, at the highest compression represented in the experiment, the syllables and words within each compressed segment of the speech record were imperfectly registered so that, when the processing time at phrase and sentence boundaries was available, some of the material to be processed was either missing or distorted.

In general, it appears that, when the time lost by compression is restored, listeners can use it to their advantage. In the language of the channel capacity model, they have more time in which to add information to long-term storage. The results of Aaronson and Markowitz (1967) also support this conclusion. Having shown this much, the next appropriate step would be to determine the amount and distribution of processing time needed to maximize comprehension at various compressions. At the time the materials for this experiment were prepared, inadequacy of equipment prevented the insertion of time at other than phrase and sentence boundaries. However, more adequate equipment is now under construction, and the research implied by the results of this experiment will soon be technically feasible.

Ideally, the question of how much time to restore to compressed speech, and at which syntactic locations, could be answered most directly by enabling the listener to control, as he listens, the restoration of time to compressed speech. The way in which he managed the restoration of time should permit inferences regarding his processing requirements. Unfortunately, with the available apparatus, it was necessary to make a decision, in advance of the experiment, about the locations at which to restore time, and about the amount of time to be restored.

There is some theoretical justification for the restoration of processing time at phrase and sentence boundaries. The units with which the listener deals may not be single words, but phrases and clauses. Foder and Bever (1965) introduced clicks at various points in sentences presented orally, and then asked their Ss where they had heard these clicks. Although the clicks were distributed throughout phrases, Ss tended to perceive their occurrence at phrase boundaries. Foder and Bever interpreted these results as providing support for the view that the segments revealed by formal constituent structure analysis function as perceptual units, and that the displacement of clicks toward phrase boundaries was an effect which insured the integrity of these units.

The hypothesis tested by this experiment is that comprehension of fluent speech depends upon processing operations which require time, and that the elimination of too much time by compression interferes with the performance of these operations. The results of this and other studies indicate that when speech occurs at a faster rate than 250 or 300 wpm, there is no longer enough time for complete processing. The silent visual reading rate typically observed may seem to argue against this conclusion. Many visual readers can read, with good comprehension, at 400 or 500 wpm, or even faster. Apparently, they can perform the needed processing operations on words put in at these rates. If they can do so when words are perceived visually, why can not they do so when these same words are perceived aurally? This dilemma is resolved by the chunking concept (Miller, 1956). As the visual reader gains experience in processing the printed page display, he organizes its units into chunks, and treats these chunks as units, thus reducing the number of units with which he must deal. Taylor (1966), in an analysis of visual reading rates, determined that the average college student makes approximately 260 visual fixations per minute. If, during each fixation, he perceived a single word, he would be reading at 260 wpm, a rate that is fairly comparable to the maximum rate at which speech can be processed aurally without loss in comprehension. However, if, through chunking, he can perceive two or more words during each fixation, he can double or even triple his visual reading rate. Because the record processed by the listener is displayed temporally, rather than visually, it is relatively inflexible, and will not yield to the kind of reorganization that can be imposed on the printed page. The listener may still try to increase the size of the units with which he deals by chunking, but his task is much more complicated. Storage and retrieval from short-term memory are added to the task of reorganizing the display. These operations must, themselves, require time, and must increase the opportunity for error.

Evidence bearing upon the foregoing analysis might be obtained by determining the maximum reading rate that could be realized by a visual reader who was compelled, through appropriate instrumentation, to perceive his display one word at a time. Under this condition, the maximum visual reading rate should approximate the maximum aural reading rate.

Summary

An experiment was performed to test the hypothesis that listening comprehension declines more rapidly than word intelligibility as a function of compression in time, because comprehension depends upon processing operations that require time. If too much time is eliminated by compression, the remaining time will not be sufficient to permit the required processing, and comprehension will suffer. The hypothesis was tested by comparing the comprehension of compressed listening selections with the comprehension of listening selections which were different only in that the time eliminated by compression was restored at phrase and sentence boundaries. The improvement in comprehension following time restoration was statistically significant for two of the three compressions at which comparisons were made. Although restoration of time at the highest compression represented in the experiment improved comprehension, it was still significantly poorer than the comprehension demonstrated by a control group that listened to uncompressed speech, suggesting that the loss of processing time is not the only factor responsible for the poor comprehension of highly compressed speech.

CHAPTER XI

SPEECH PAUSE DURATION AS A FUNCTION OF SYNTACTIC JUNCTURES

Kenneth F. Ruder and Paul J. Jensen*

Introduction

The phenomenon of speech is often associated with images which suggest continuity in sound production; i.e., we speak of fluency in speech, the flow of speech, etc., yet evidence shows that at least 40-50% of speaking time is spent in pausing (Goldman-Eisler, 1968). This fact, then, shows that the act of speaking is anything but continuous. It is fragmented and interrupted throughout by intervals of silence--that is, by pauses.

Despite the frequency of occurrence of speech pauses, most of the research in the production and perception of speech has been concerned with the filled elements of utterances, that is, vocal sounds that are intended to carry linguistic information. In fact, it has only been within the last decade that researchers (Boomer & Dittman, 1962; Goldman-Eisler, 1968; Maclay & Osgood, 1959) have become seriously interested in the unfilled events of utterances--speech pauses.

As a result of this general disregard, actually very little is understood about the nature and function of pauses in speech. A number of speculations have been advanced in this regard, but for the most part these speculations have not been subjected to systematic and rigorously controlled research until very recently.

One such speculation concerns the extent to which speech pauses may be related to encoding and decoding decisions concerning sentence structure. In general speech literature, pauses are often referred to as oral punctuation. Implicit in such a report is the notion that somehow pauses in speech reflect and signal the syntactic structure of the sentence. However, such a relationship between speech pauses and constituent structure is at present nothing more than a broad sweeping generalization which has yet to be verified.

---

*Dr. Kenneth F. Ruder is affiliated with the Bureau of Child Research, University of Kansas, Lawrence, Kansas 66044. Dr. Paul J. Jensen is with the Communication Science Laboratory, University of Florida, Gainesville, Florida 32601.

Lieberman (1967) does provide some data which shows that speech pauses do serve as disjuncture cues which reflect the constituent structure of speech. His data, in this regard, are based on an experiment in which computerized tape splicing techniques were used to change the utterance "lighthouse keeper" into "light housekeeper" and vice versa by simply changing the interval between light and house. Lieberman (1967) does state, however, that disjunctures would manifest the constituent structure only where it would not otherwise be clear from the total context of the message.

The role of pauses in signaling the constituent structure of sentences in normal speech is, thus, still not clarified since in the context of normal speech the redundancies of language minimize the occurrences of ambiguous utterances such as "light house keeper." Scholes (1968), in fact, refutes Lieberman's (1967) claim that the speech pause is a significant cue for disjuncture on the basis of an acoustic analysis of disambiguating disjunctures occurring in normal speech.

It should be pointed out that in both the Lieberman (1967) and Scholes (1968) studies, the disjuncture cues were studied within the context of an ambiguous utterance--a not too frequent occurrence in everyday speech. Wilkes and Kennedy (1969) studied the distribution of pause durations as a function of constituent structure in normal, non-ambiguous utterances and found that the subject-predicate break, relative clauses, and enumerated items ("cats goats pigs ducks make good pets") were marked by relatively longer pauses than the rest of the sentence constituents. This data could thus be interpreted as indicating that the function of speech pauses is, indeed, to signal the syntactic structure of the sentence.

Up to this point we have talked only about pauses occurring in a fluent stretch of speech--fluent pauses, if you will. There are, however, pauses in speech which perceptually disrupt the smooth flow of speech. These pauses are referred to in the literature as hesitation pauses. Goldman-Eisler (1968) reports on the distribution of hesitation pause durations in reading and spontaneous speech and reports that the constituent structure, no matter how complex, was not reflected in hesitation pause duration.

These results would seem to be in direct opposition to the Wilkes and Kennedy (1969) data. However, it may merely reflect the notion that the function of the two pause types is different--a not too novel interpretation. Since the pause measures in the two studies were obtained in completely different contexts, the two studies cannot realistically be compared. To resolve this dilemma to some extent, the question can be posed as to whether fluent and hesitation pauses differ with respect to reflecting constituent structure (as the two aforementioned studies indicated) and to what extent does either type of pause adequately reflect sentence structure.

In regard to these considerations, then, the following study was undertaken to investigate the duration of fluent and hesitation pauses as a function of syntactic complexity.

Procedure

Five stimulus sentences were chosen in which the words "lost" and "contact" were manipulated so that they assumed different syntactic relations to one another. These syntactic relations were specified by means of tree diagrams which defined the constituent structure of each stimulus sentence.

Since the purpose of this study was to look at pause duration as a function of syntactic complexity, it was necessary to quantify the complexity of the syntactic relations between "lost" and "contact" so that these relations which constitute the syntactic environment could be rank ordered. This quantification was accomplished through use of structural complexity index (Miller & Chomsky, 1963).

To apply this metric, one locates the lowest node of the tree diagram which encompasses both words of interest, "lost" and "contact," and counts the total number of nodes contained in this branch of the tree structure. This number is then divided by the number of terminal nodes contained in that branch (see Figure 11.1). For instance, in the sentence "Even though the battle was lost, contact had been established with the rear guard.", the lowest node which encompasses the words "lost" and "contact" is the sentence node. The total number of nodes in this branch is 27. The number of terminal nodes is 14. The node-to-terminal-node ratio of 27/14 (1.93) is thus the structural complexity index for that particular syntactic relation between the words "lost" and "contact." Since the constituent structure analysis used in arriving at the tree diagram is an arbitrary procedure in many instances, an external test of the validity of the syntactic complexity rankings based on the node-to-terminal-node ratio was made. Eighteen naive Ss were asked to rank the stimulus sentences on the basis of the complexity of the syntactic relations between the words "lost" and "contact." Subjects were merely told to use their own judgment in regard to what was meant by syntactic relations and complexity and were given no special instructions or training in this regard.

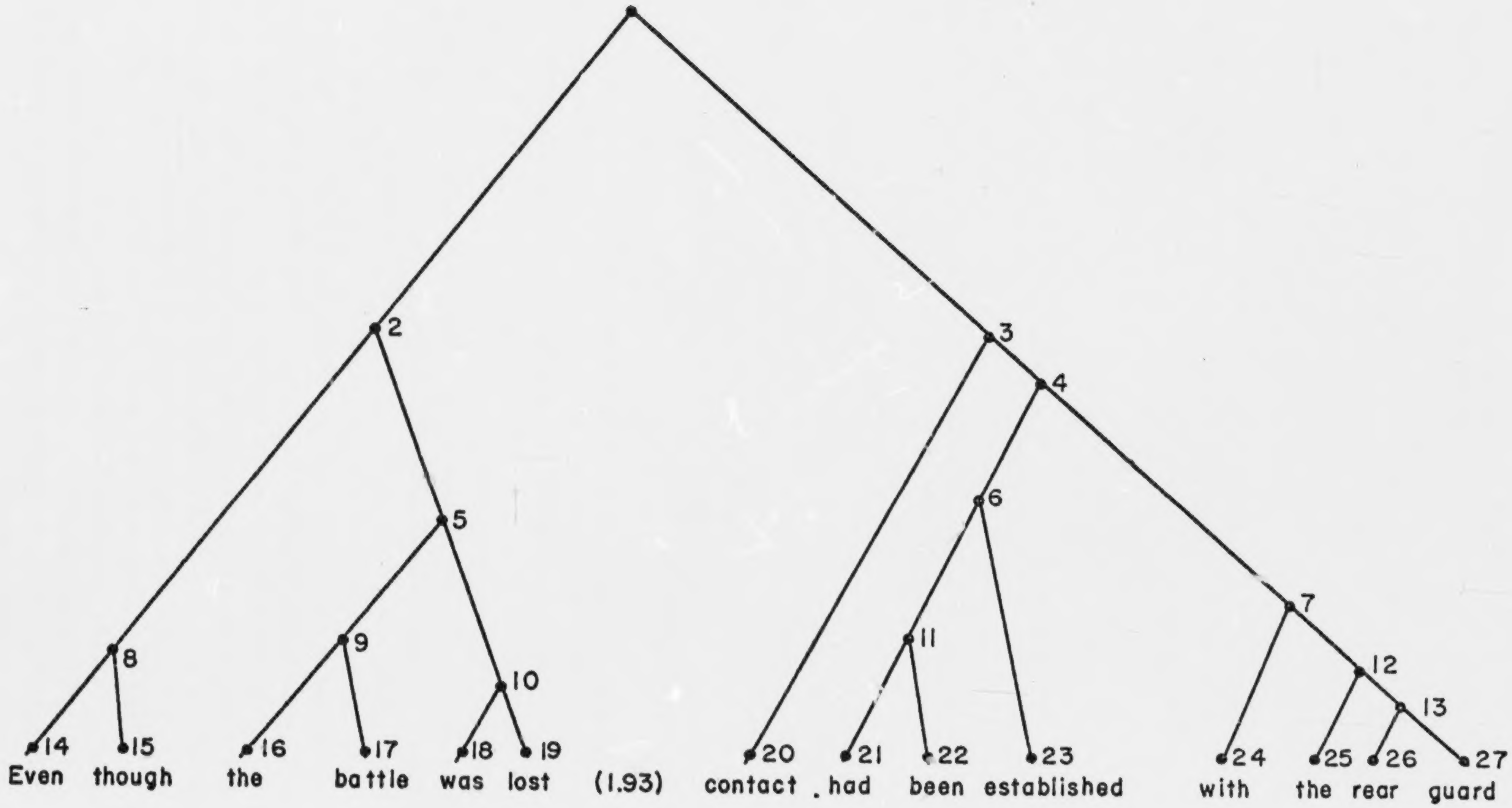

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
Even though the battle was lost (1.93) contact . had been established with the rear guard

Figure 11.1. Tree diagram analysis of stimulus sentence five. The node-to-terminal-node ratio specifying the complexity of the relation between "lost" and "contact" is 27/14 (1.93).

The Spearman Rank Order Correlation between the mean complexity rankings of the Ss and the complexity rankings based on the node-to-terminal-node ratio was .96. There is, therefore, a high degree of relationship between naive Ss' concept of syntactic complexity and the structural complexity index based on the node-to-terminal-node ratio. Insofar as the naive Ss' complexity rankings are viewed as adequate criteria, the structural complexity index as derived from an immediate constituent analysis thus seems to be a valid measure of the complexity of syntactic relations.

The stimulus sentences used in the study were, in order of increasing complexity of the syntactic relations between "lost" and "contact":

1. He lost contact with reality.
2. The lost contact lens case was finally found.
3. The team that lost contacted the authorities.
4. If you are lost, contact the nearest policeman.
5. Even though the battle was lost, contact had been established with the rear guard.

The psychophysical method of adjustment was employed to obtain the perceptual judgments of pause duration between the words "lost" and "contact" in each of the five stimulus sentences. Adjustment of pause duration between these two words was accomplished in the following manner. The five stimulus sentences were recorded by a single speaker who had been selected from an original pool of four on the basis of his high ranking on the six speaker attributes of quality, fluency, naturalness, rate, precision, and overall effectiveness.

These sentences were then dubbed onto a two-track tape such that the portion of the sentence up to and including the word "lost" was recorded on Track A, while the remainder of the sentence (beginning with the word "contact") was recorded on Track B. Pause duration between the words "lost" and "contact" could then be manipulated by mechanically delaying, to a greater or lesser degree, the playback of the Track B signal with respect to the Track A signal.

Experimental Apparatus

To implement this procedure, a modified two-track playback head system was constructed so that pause duration adjustment could be accomplished through delaying, to a greater or lesser degree, the playback of the Track B signal (that part of the sentence beginning with the word "contact") relative to the signal on Track A (that part of the sentence ending with the word "lost"). To provide for such a variable delay, a tape guide attached to a worm gear assembly, as shown in Figure 11.2, was mounted between the fixed Track A and Track B playback heads. Since

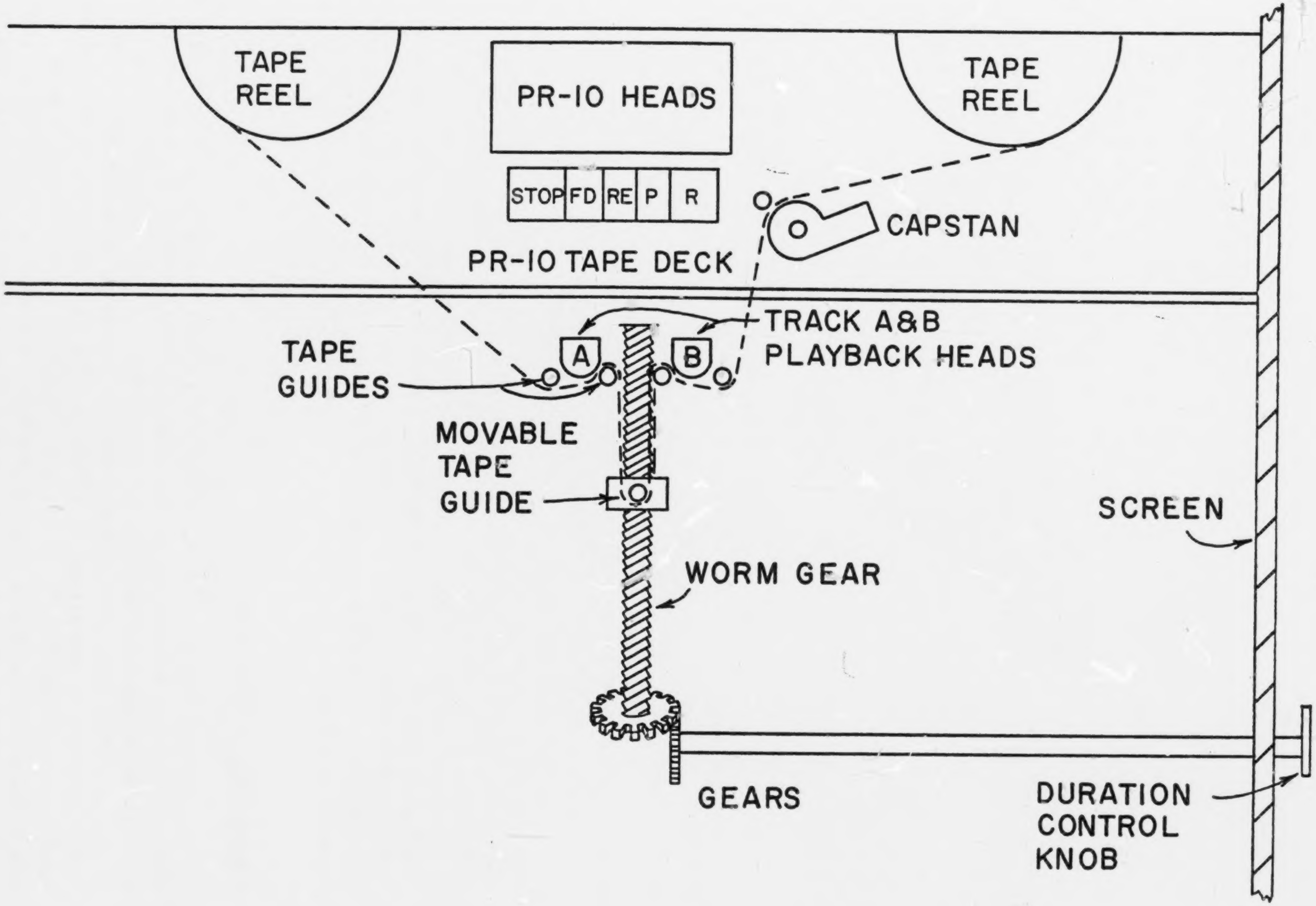

TAPE REEL
PR-10 HEADS
TAPE REEL
STOP
FD
RE
P
R
CAPSTAN
PR-10 TAPE DECK
TRACK A&B PLAYBACK HEADS
A
B
TAPE GUIDES
MOVABLE TAPE GUIDE
WORM GEAR
SCREEN
GEARS
DURATION CONTROL KNOB

Figure 11.2. Instrumentation for variable delay system to adjust pause durations.

the tape guide was attached to the worm gear, its position, relative to the horizontal plane of the playback heads, could be varied up and down the worm gear, thereby increasing or decreasing the effective distance between the playback heads. Pause duration between "lost" and "contact" could be increased, then, by moving the tape guide away from the playback heads, since this movable guide altered the time of onset of the Track B signal relative to that on Track A by changing the distance the tape must travel to reach the "B" playback head.

This playback head assembly was coupled to an Ampex PR-10 tape deck and stereo amplifier assembly driving a pair of Sharpe HA-10 earphones with a monaural Y-cord input so that the Track A and B signal was received in both ears (see Figure 11.3).

The outputs from Channel A and Channel B of the playback recorder were coupled respectively to Channel A and B inputs of a Tektronix 564 storage oscilloscope. This oscilloscope is a special purpose instrument designed to store cathode-ray tube displays for viewing or photographing. Each stimulus sentence, as adjusted by the S, was viewed on the oscilloscope screen. When that portion of the signal containing the end of Channel A signal and beginning of Channel B signal was located visually, it was stored on the oscilloscope display. A permanent record of this display was then obtained by photographing it with a Polaroid trace-recording camera attached to the face of the oscilloscope screen. Pause durations were obtained from these photographs by measuring the time interval from the cessation of the Track A signal to the beginning of the Track B signal (see Figure 11.4).

Twelve young adult male college students were selected as Ss for the study. Each S, working individually, was asked to adjust the duration of the pause between the words "lost" and "contact" so that (1) the pause duration was considered optimal for the "fluent" presentation of the sentence, and (2) the pause was just perceived as being a hesitation.

To provide some data concerning the range of fluent pauses (that range of durations extending from the pause detection threshold in speech to that point where the pause was just perceived as a hesitation), each S's pause detection threshold, at each level of syntactic complexity, was found.

In this detection task, the Ss were asked to adjust the duration of the silent interval between the words "lost" and "contact" to that point at which they just detected a pause. This adjustment began with the pause duration between "lost" and "contact" set at 500 milliseconds (msec.)(arbitrary starting point). The S was instructed to decrease the duration of this pause until he perceived a definite overlap of the words "lost" and "contact" and then increase the duration again until he just detected a pause. The S was encouraged to make as many such "threshold crossings" as necessary in order to locate that point at which the duration of the silent interval was just sufficient to be perceived as a pause, that is, if the

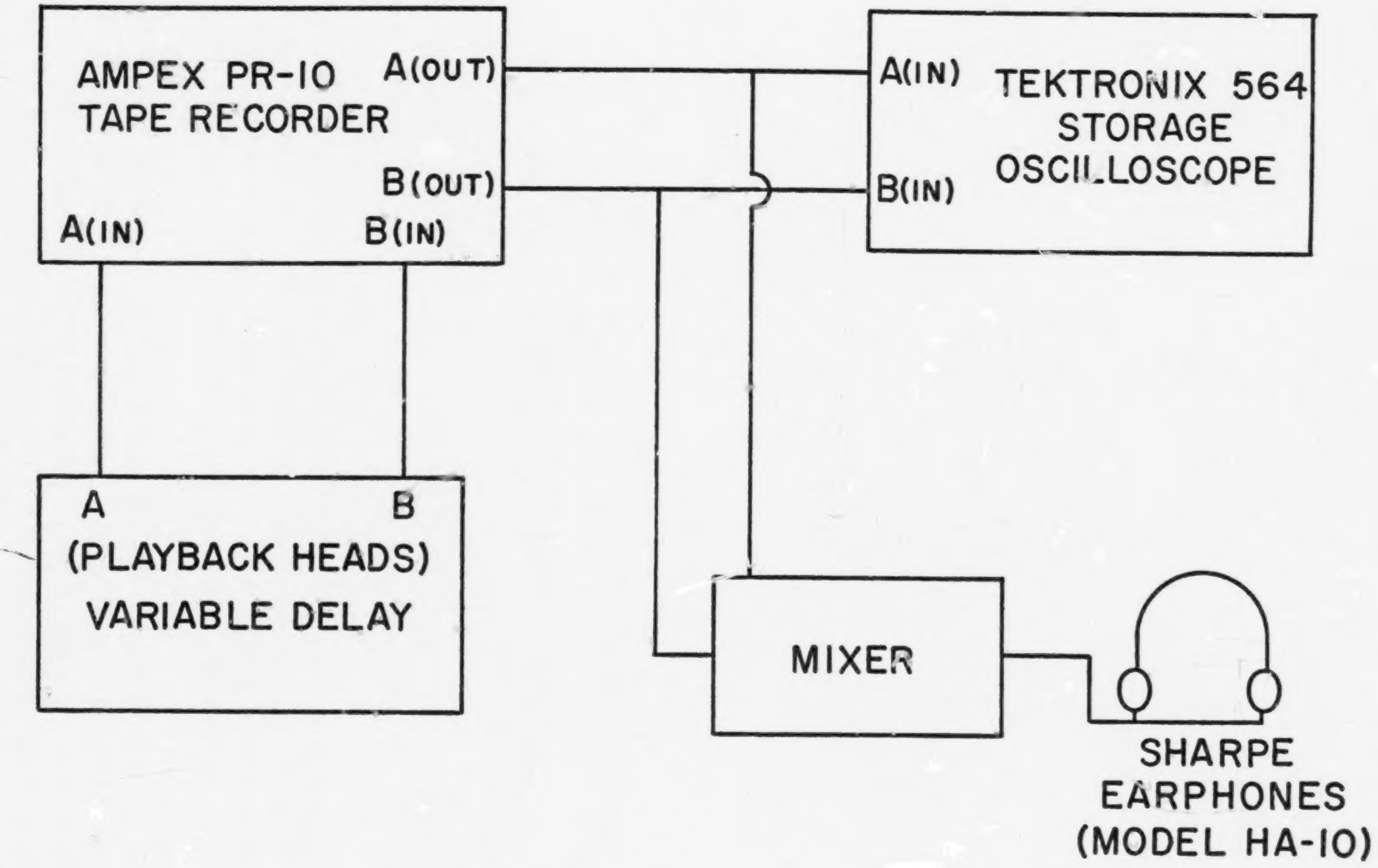

AMPEX PR-10
TAPE RECORDER
A(OUT)
B(OUT)
A(IN)
B(IN)
TEKTRONIX 564
STORAGE
OSCILLOSCOPE
A(IN)
B(IN)
A
B
(PLAYBACK HEADS)
VARIABLE DELAY
MIXER
SHARPE
EARPHONES
(MODEL HA-10)

Figure 11.3. Block diagram of instrumentation for experimental pause adjustment and measurement.

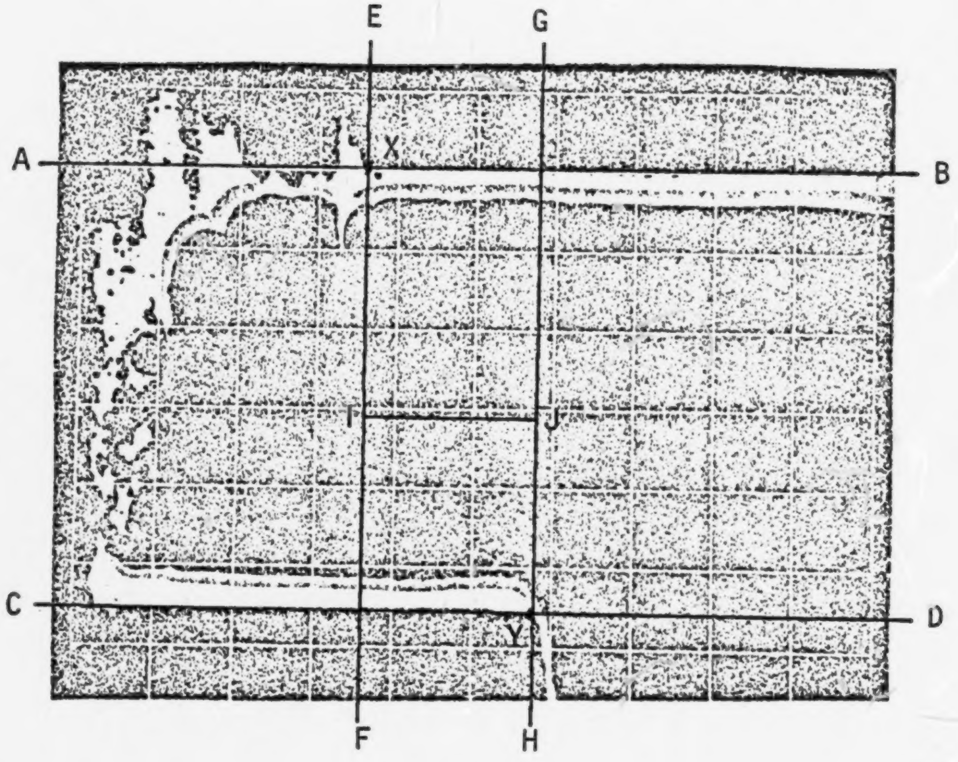


Figure 11.4. Reference points used in the measurement of pause durations from the oscillographic photographs: line AB, the top of the Channel A baseline; line EF, a perpendicular erected through AB at point X (point where the speech signal drops to the baseline AB); line CD, the bottom of the Channel B baseline; line GH, a perpendicular erected at point Y (point at which the Channel B integrated signal drops below the baseline); line IJ, the distance in centimeters from the end of Channel A signal to the beginning of Channel B signal.

duration were any shorter, the interval would not be perceived as a pause. When the S indicated he had located this point, the adjusted sentence was displayed on the oscilloscope screen and photographed for subsequent analysis.

For the task which required adjustment of the optimal duration of fluent pauses, as well as for the adjustment of hesitation pauses, the S was merely informed that there are two types of pauses: (1) pauses which do interrupt the smooth flow of speech (hesitation pauses), and (2) pauses which do not interrupt the smooth flow of speech (fluent pauses). No more detailed description of these pauses was given (such as: fluent pauses are said to be used as oral punctuation, or, hesitation pauses are silent intervals of unusual length), since such specific information may have inadvertently influenced the Ss' performances.

For the fluent pause task, the Ss were presented a sentence in which the duration of the silent interval to be adjusted was arbitrarily set at 500 msec. They then were instructed to adjust the duration of this interval, either increasing or decreasing it, to the point where it was perceived as being an optimal duration, that is, the duration which was best suited for conveying the meaning of the sentence in a fluent, intelligible manner. Subjects were allowed as many trials as necessary in making this adjustment; when they indicated that the duration of the pause was optimal for the fluent presentation of the sentence, it was displayed on the oscilloscope screen and photographed.

The hesitation pause task proceeded exactly as the fluent pause task except that Ss were requested to adjust the pause duration to that point where they just perceived it as being a hesitation. Each pause adjustment for the five sentences within each task was repeated three times to obtain an estimate of within S variability. Subjects were given as much time and as many trials as necessary to make each adjustment.

Results and Discussion

The overall mean pause duration for each level of syntactic complexity is shown in Figure 11.5. A treatments by Ss (Winer, 1962) analysis of variance indicated that the main effect of syntactic complexity was statistically significant. However, there was also a significant interaction between tasks and levels of syntactic complexity; therefore, a simple main effect test was performed for syntactic complexity at each level of tasks.

The results of this analysis indicated that level of syntactic complexity was a significant factor only for the hesitation pause task. Figure 11.6 clearly illustrates that as the level of syntactic complexity increases

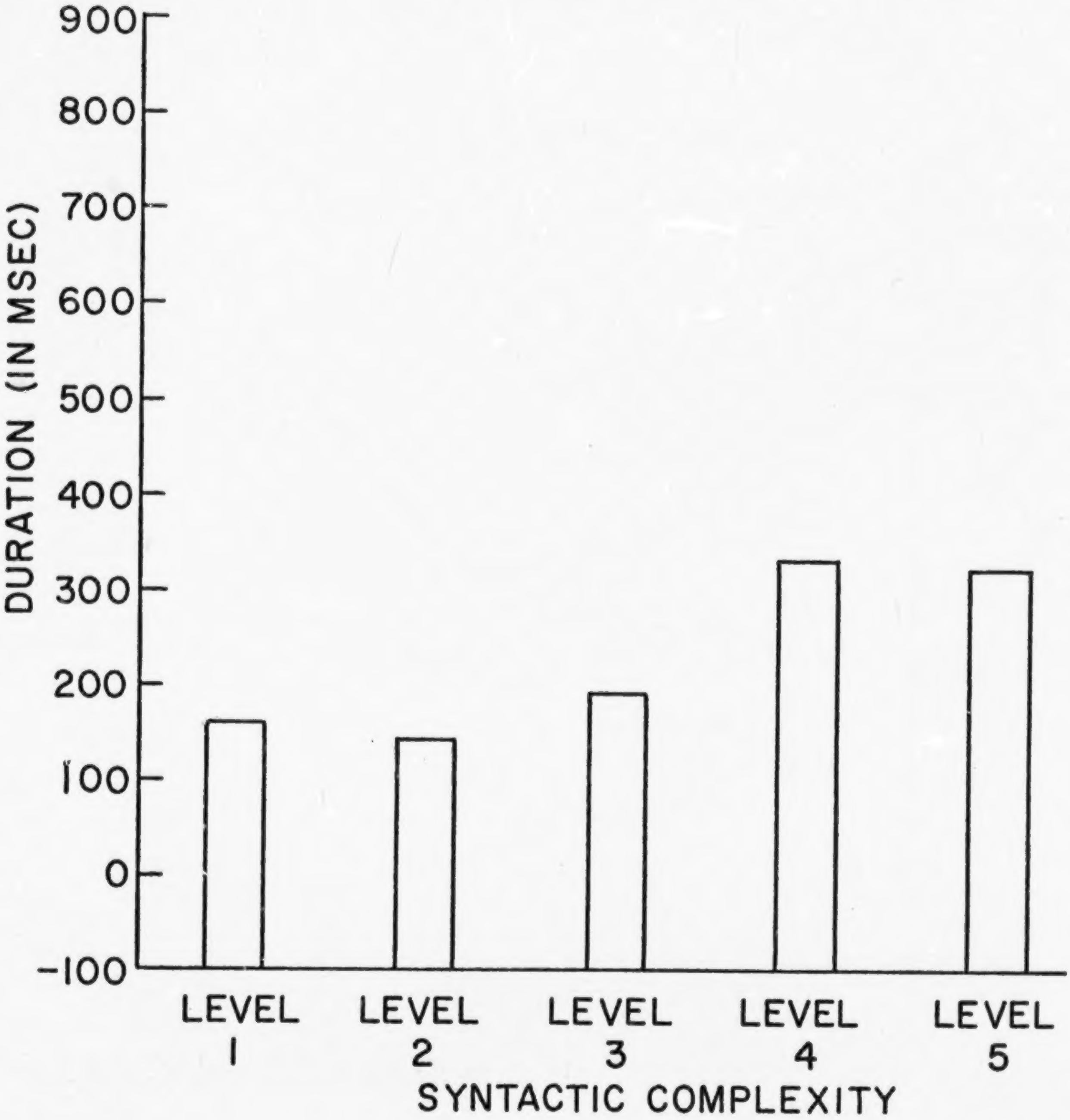


Figure 11.5. Mean pause duration for each of five levels of increasing syntactic complexity.

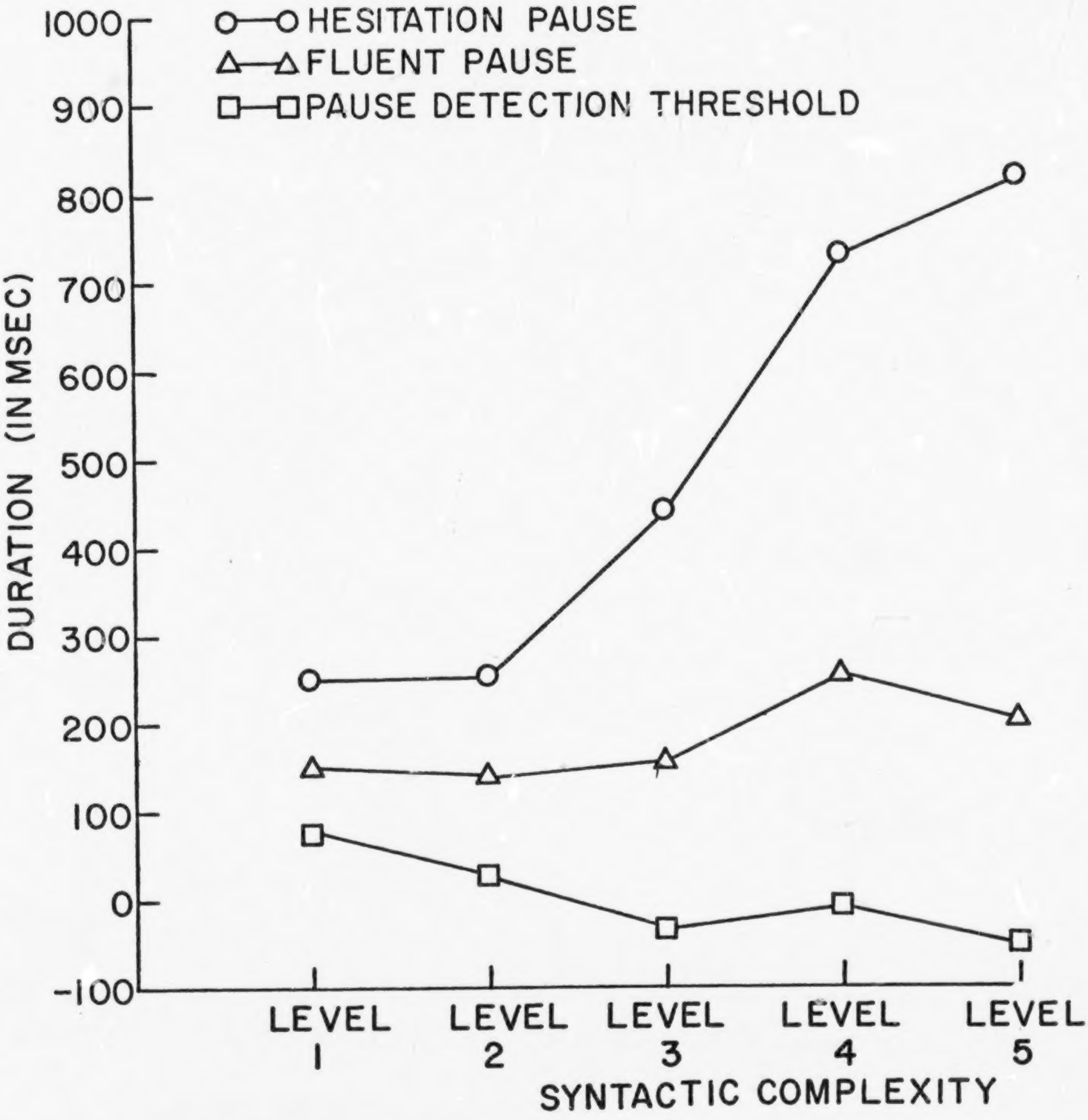


Figure 11.6. Mean pause duration of the five levels of increasing syntactic complexity for each task.

from levels two to five, the duration of the hesitation pause similarly increases. The increase in hesitation pause duration from syntactic complexity level one to level two, however, is slight.

To determine which of these levels of complexity was contributing to the overall simple main effect, a Newman-Keuls test of treatment differences was applied to the data. The results of this analysis indicate that mean hesitation pause durations at levels four and five differ significantly from the mean hesitation pause durations at levels one, two, and three. Hesitation pause durations at levels four and five, however, do not differ significantly, nor are there any significant differences in the durations among levels one, two, and three.

These results can be interpreted as suggesting that syntactic complexity as measured by the two indices, structural complexity index and Ss' rankings of syntactic complexity, does not have an effect on pause durations. While the syntactic complexity of the pause boundaries in sentences four and five differed considerably on both complexity indices, hesitation pause duration did not differ significantly at these boundaries. Sentences one, two, and three are similarly different from each other in terms of the complexity indices and likewise show no significant differences in hesitation pause duration among themselves. In this respect, the pause boundaries in sentences four and five are alike in that both are subordinate clause-main clause boundaries, while those in sentences one, two, and three are other types of syntactic boundaries (i.e., verb-object boundaries, adjective-noun boundaries, etc.). For convenience, the former will henceforth be referred to as subordinate clause boundaries and the latter simply as other boundaries.

Fluent Pause Duration as a Function of Syntactic Complexity

The finding that fluent pause duration did not differ significantly across the five levels of syntactic complexity has import in regard to the assumption that fluent pauses are used to signal the syntactic structure in speech (Carrell & Tiffany, 1960; Lieberman, 1967). On the basis of these results, one would have to conclude that this is not one of the primary functions of fluent pauses. This has particular relevance to the notion advanced by Lieberman (1967) and Wilkes and Kennedy (1969) that fluent pause duration is a primary cue marking syntactic junctures in speech. If this notion were, in general, correct, one would expect a difference in fluent pause durations at least for the between-phrase and within-phrase distinctions. Since such was not the case, the data from the present study are seen as contradictory evidence for the claim that fluent pause duration is a primary cue reflecting constituent structure. On the other hand, the data can be interpreted as direct support for the interpretation advanced by Scholes (1968) that fluent pause duration is a relatively unimportant physical cue for signaling syntactic junctures (disjunctures).

This is not to say that fluent pause duration may not be used as a cue for disjuncture. The data merely suggest that, as a general rule, duration of the silent interval is not utilized for this purpose. It may also well be that given specific instructions to emphasize or de-emphasize a particular part of the sentence, pause duration will be a primary parameter employed. Such a function of fluent pauses is suggested by Carrell and Tiffany (1960).

Then again, fluent pause duration may serve no other purpose in speech than to mark perceptual units in speech and to provide perceptual processing time as suggested by Aaronson and Markowitz (1967). These are only speculations, however, and must be researched further. The data from the present study offer little insight with regard to these possible functions of fluent pauses.

Hesitation Pause as a Function of Syntactic Complexity

A comparison among mean hesitation pause durations at the five levels of syntactic complexity shows clearly that the most complex syntactic relations (levels four and five) are associated with significantly longer hesitation pause durations than are the less complex relations (levels one, two, and three). As mentioned previously, the variable pause boundaries in sentences four and five differ from those in sentences one through three in that the former are subordinate clause boundaries while the latter are other types of syntactic boundaries. Since subordinate clause boundaries are frequently punctuated by a comma in orthography, these results can be viewed as being particularly relevant to the notion that fluent pauses in speech are often used as oral punctuation (Carrell & Tiffany, 1960). Although the comparison of fluent pause durations did not reveal a significant differential use of pause durations between subordinate clause and other syntactic boundaries, the use of hesitation pause durations in this regard has some interesting implications.

The major implication of this finding is that although the optimal fluent pause duration between subordinate clause and other boundaries does not differ significantly, the range of fluent pause durations (that range of durations extending from the pause detection threshold to the minimum hesitation pause duration) is greater at the subordinate clause boundaries (917 msec.) than at the other boundaries (496 msec.). Although this range may not be utilized extensively, its potential is there as can be evidenced by this difference.

An equally important finding is that the within-phrase, between-phrase distinction frequently employed in the study of hesitation pauses

(Goldman-Eisler, 1968; Maclay & Osgood, 1959) may not be as significant as commonly thought. Goldman-Eisler (1968), in particular, frequently limits her study of hesitation pauses to within-phrase occurrences such as those corresponding to syntactic complexity levels one and two of this study. In this regard, it may be interesting to note that the mean within-phrase pause duration in the present study of 265 msec. compares favorably with the 250 msec. standard adopted by Goldman-Eisler (1961a) and to a lesser degree, the 200 msec. "hesitation pause threshold" reported by Boomer and Dittman (1962) at within-phrase boundaries. However, the data in the present study also indicate that the hesitation pause durations at the syntactic complexity level three are not significantly different from those at levels one and two, yet the former is a between-phrase boundary and the latter are within-phrase pause boundaries. This suggests that any generalization about hesitation pauses based largely on within-phrase occurrences is of questionable validity. Furthermore, the differences between hesitation pause durations at subordinate clause boundaries and other boundaries argue rather forcibly against the notion that a hesitation pause threshold can be established irrespective of environment and/or perceptual judgments as has been commonly assumed (Boomer & Dittman, 1962; Goldman-Eisler, 1961a, 1968).

The distinctions of within-phrase and between-phrase pauses and the subordinate clause and non-subordinate clause pauses may, in this regard, be further useful to demonstrate the behavioral relevance of these units. That is, the differential distribution of hesitation pauses at the subordinate clause and other types of syntactic boundaries can be interpreted as evidence supporting the functional distinction between such boundaries.

ACKNOWLEDGMENTS

This research was supported by NSF Grant GS-2337 and by Grant NB-06459 of the National Institute of Neurological Diseases and Stroke and was part of a doctoral dissertation completed under the direction of Dr. Paul Jensen.

CHAPTER XII

THE SIGNIFICANCE OF INTRA- AND INTER-SENTENCE PAUSE TIMES IN PERCEPTUAL JUDGMENTS OF ORAL READING RATE*

Norman J. Lass**

Normally when one attempts to alter his oral reading rate, he exhibits changes in both speech and pause times (Lass, 1968; Minifie, 1963). However, several investigators have emphasized the importance of pauses and pause times in the determination of reading rate as well as in the perceptual judgment of reading rate (Agnello, 1963; Franke, 1939; Minifie, 1963). Since pause time is regarded so highly in perceptual judgments of rate, an investigation was undertaken to determine if alterations in intra- and inter-sentence pause times in an individual's recorded reading of a standard prose passage (with no changes made in speech time) would result in alterations in perceptual judgments of his oral reading rate.

Procedure

Reading Material

The "Rainbow Passage" (Fairbanks, 1960) was employed for pause alteration purposes in the study. The passage was read by a 30-year-old male with judged normal voice, articulation, and rate characteristics. Recording of the reading was made in an IAC chamber using an Electro-Voice model 664 dynamic cardiod microphone and an Ampex model 602 tape recorder.

---

*To be published in Journal of Speech and Hearing Research in 1970.

**Dr. Norman J. Lass is with the Speech and Hearing Sciences Laboratory, Division of Otolaryngology, School of Medicine, West Virginia University, Morgantown, West Virginia 26506.

Location of Pauses

A group of 20 individuals determined the location of pauses in the reading by listening to the recording and placing pencil marks on transcribed copies of the reading where they thought pauses occurred. It was necessary for 15 of the 20 listeners (75%) to agree on a pause point in order for it to be considered and used as a pause in the study. In this manner, a total of five intra-sentence and five inter-sentence pauses were located in the reading.

Pause Alteration Procedure

Duration of intra- and inter-sentence pause times was determined by means of an Ampex model 602 tape recorder and a Bruel and Kjaer model 2305 power level recorder. It was found that the intra-sentence pause time for the entire reading passage was 2.2 seconds, while duration of inter-sentence pause time totaled 3.5 seconds.

It was arbitrarily decided to manipulate intra- and inter-sentence pause times in one of seven ways: (a) no change, (b) 25% increase, (c) 25% decrease, (d) 50% increase, (e) 50% decrease, (f) 75% increase, and (g) 75% decrease. Any changes of pause times involved either increasing both pause types or decreasing both of them. Increases in one pause type and decreases in the other were not used. A total of 31 pause alteration conditions were employed (see Table 12.1). Thirty electronic reproductions of the original recording were made; the removal or addition of pause time was accomplished in all cases by the removal or addition of the appropriate percentages of actual tape at each pause point. All pause alterations in a particular recording were distributed approximately equally over the total number of pauses involved. Thus, for example, if a total of 5.0 seconds were to be added to (or subtracted from) the intra-sentence pause time in a recording, each of the five intra-sentence pauses would be increased (or decreased) by 1.0 second. To control for background noise in all recordings, in those recordings where tape was to be added, tape added was erased on the same recorder used for the original recording.

After all pause alterations were completed, the 30 altered recordings, plus the original unaltered recording, were arranged in random order on an experimental master tape. In addition, five of the 31 recordings (15%) were randomly selected and included at the end of the master tape for intra-judge reliability estimation. Thus, a total of 36 readings were included on the master tape.

TABLE 12.1

MEAN RATINGS, PAUSE ALTERATION CONDITIONS, AND OVERALL READING TIME (IN SECONDS) FOR THE 31 READINGS OF THE RAINBOW PASSAGE

	Pause Alterations		
Mean Rating	Intra-Sentence	Inter-Sentence	Time (sec.)
2.6	+75%	+75%	31.4
2.6	+50%	+75%	30.9
2.8	+75%	+50%	30.8
2.6	+25%	+75%	30.7
2.8	+50%	+50%	30.3
2.8	0	+75%	30.2
3.1	+75%	+25%	30.2
3.1	+25%	+50%	29.8
3.2	+50%	+25%	29.7
3.1	+75%	0	29.6
3.0	0	+50%	29.5
3.2	+25%	+25%	29.2
3.3	+50%	0	29.0
3.5	0	+25%	28.9
3.5	+25%	0	28.8
3.4	0	0	28.2
3.8	-25%	0	27.7
4.0	0	-25%	27.3
3.9	-50%	0	27.2
3.9	-75%	0	26.9
4.0	-25%	-25%	26.7
4.1	0	-50%	26.6
4.0	-50%	-25%	26.4
4.3	-25%	-50%	26.1
4.5	0	-75%	26.0
4.3	-75%	-25%	25.9
4.4	-50%	-50%	25.6
5.1	-25%	-75%	25.4
4.6	-75%	-50%	25.3
4.9	-50%	-75%	25.0
5.2	-75%	-75%	24.7

Rating Session

Judges. A total of 78 individuals, 40 males and 38 females, served as judges. All were volunteer Ss from a basic psychology course at the

University of Kansas. The group ranged in age from 18 to 28 years, with a mean age of 19 years.

Rating scale. A 6-point equal-appearing interval scale was employed for evaluation of reading rate (Agnello, 1963). On the scale, the numbers represented the following rates:

1 = very slow
2 = slow
3 = moderate
4 = moderate
5 = fast
6 = very fast

The judges were asked to listen carefully to the entire reading before circling the number which they thought best represented the reading rate for that particular reading.

Instructions to judges. A set of prerecorded instructions was included on the master tape. It included an explanation of the 6-point equal-appearing interval scale, examples of slow, moderate, and fast reading rates, and five readings to be evaluated by the judges for practice purposes. All examples and practice readings involved the same reader and reading passage as was used in the experimental recordings. Upon completion of the instructions, all questions of the judges were answered and the experimental readings were played. Playback equipment included an Ampex model 602 tape recorder and an Ampex model 622 speaker-amplifier system.

The rating session, which lasted approximately 60 minutes, was held in a quiet room in one of the buildings on the University of Kansas campus.

Results

Table 12.1 contains the mean ratings of the judges for each of the 31 pause-altered readings as well as the total time, in seconds, for each reading.

The results indicate that changes in intra- and inter-sentence pause times appeared to influence changes in the perceptual judgments of oral reading rate. In addition, the resulting perceptual changes were in the expected directions. That is, with increases in pause times, there were corresponding changes toward lower ratings on the scale (slower perceived rates); decreases in pause times produced judgmental changes toward higher ratings (faster perceived rates). Thus, it appears that there is a strong inverse relationship between pause times and perceptual judgments of oral reading rate. However, as can be seen from Figure 12.1, the overall reading time (in seconds) also appears to be related to perceptual

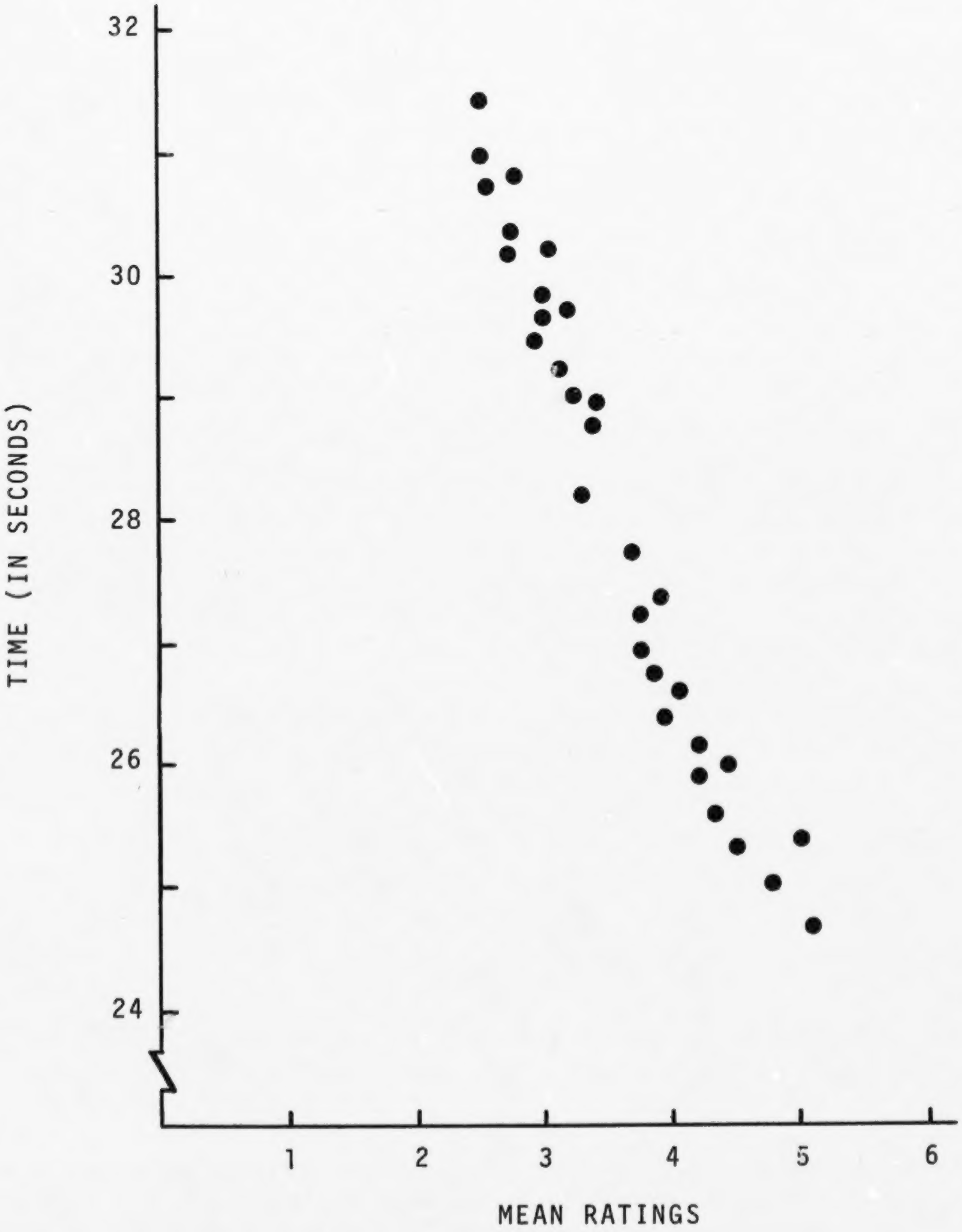


Figure 12.1. A scattergram displaying the relationship of overall reading time (in seconds) to mean ratings of rate.

judgments of rate (r = -.99). Since it was impossible to alter pause times without also altering overall reading times, it was necessary to determine if the changes in mean ratings were due to changes in overall reading time alone or also to pause time alterations. An analysis of covariance, involving adjustment for a linear dependence of ratings on time (Williams, 1959; Winer, 1962), was employed to answer this question. The obtained F value, which was significant at the .01 level, indicated that pause time alterations had a significant effect on changes in mean ratings.

Intra- versus Inter-sentence Pauses

Figure 12.2 shows a comparison of intra- and inter-sentence pause time alterations and their effect on mean ratings of reading rate. It appears that inter-sentence pause time alterations affected mean rate ratings more than did alterations of intra-sentence pause time. The range of mean ratings for inter-sentence pause time alterations was 2.8 (slow rate) to 4.5 (moderate rate), while intra-sentence pause time alterations showed a range of mean ratings of 3.1 (moderate rate) to 3.9 (moderate rate).

In addition, in those pause alteration conditions where both intra- and inter-sentence pause times are altered by equal percentages, inter-sentence pause time alterations show greater changes in mean ratings than those involving intra-sentence pause time. For example, in Table 12.1, a 75% increase in intra-sentence pause time and a 25% increase in inter-sentence pause time produced a mean rating of 3.1 (moderate rate); while a 25% increase in intra-sentence pause time and a 75% increase in inter-sentence pause time produced a mean rating of 2.6 (slow rate). However, since in the original unaltered recording, total inter-sentence pause time is greater than total intra-sentence pause time, any changes in inter-sentence pause time will result in longer overall reading time than identical percentage changes in intra-sentence pause time. Therefore, from the results of this study, it is impossible to make any definitive statement concerning the relative importance of these two types of pauses in influencing perceptual judgments of reading rate.

Intra-judge Reliability

To obtain an estimate of intra-judge reliability, mean discrepancy scores were computed for each of the five repeated readings on the master tape. A grand mean discrepancy score of 0.62 of a scale value was obtained, indicating a satisfactorily small dispersion within each of the 78 judges between his first and second ratings of the five repeated readings.

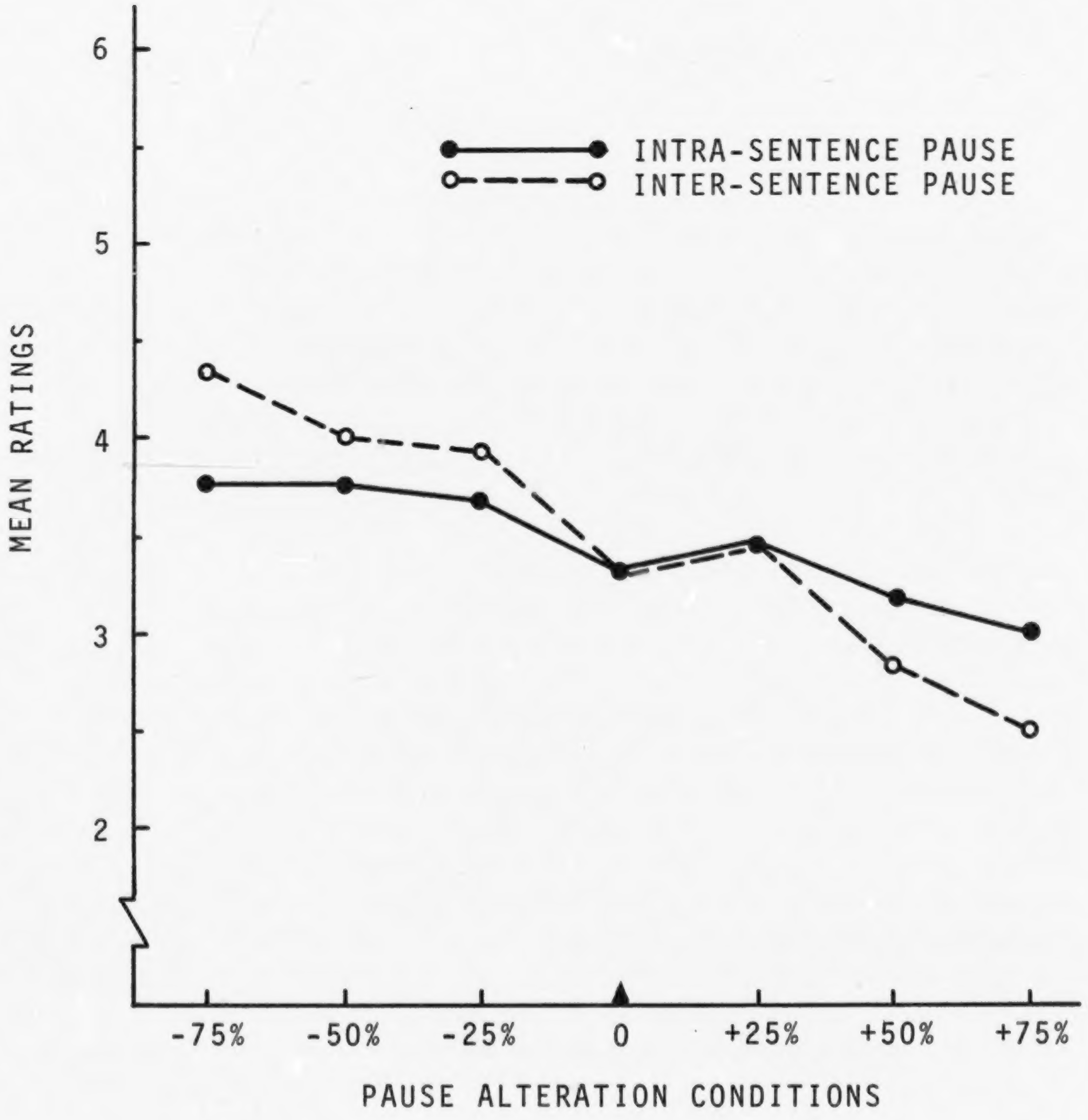


Figure 12.2. Graph showing the effects of intra- and inter-sentence pause time alterations on mean ratings of reading rate.

Inter-judge Reliability

Inter-judge reliability was estimated by means of an analysis of variance (Winer, 1962). An r of .99 was obtained from this analysis. The interpretation of this finding is as follows: if the experiment were to be repeated with another sample of 78 judges, but with the same readings, the correlation between the mean ratings obtained from the two sets of data on the same readings would be approximately .99.

Discussion

The results indicate that pause time appears to play a definite role in affecting perceptual judgments of oral reading rate. It should be noted that speech time was in no way altered; it remained the same for all of the 31 readings. Moreover, the alterations in pause time were relatively small. A 75% change, the largest amount allowable in the study, involved only a change of approximately 1.7 seconds for intra-sentence pause time and 2.6 seconds for inter-sentence pause time. The largest time change possible for combinations of intra- and inter-sentence pause time alterations (75% change in both intra- and inter-sentence pause times) totaled only approximately 4.2 seconds. Nevertheless, such pause changes appear to have affected the listeners' perceptual judgments of reading rate. It is assumed from these findings that larger changes in pause times than the ones used in this study would produce greater changes in perceptual judgments of rate.

The results of this investigation may provide some very useful clinical information for speech pathologists. Since alterations in pause time, with speech time unaltered, have produced changes in perceptual judgments of rate, perhaps emphasis in speech therapy on pause time alterations, rather than the traditional emphasis on speech time alterations (Fairbanks, 1960; Weiss, 1964), would result in the desired changes in perceptual judgments of rate. In addition, since it has been found that when one attempts to change his reading rate in a given direction (increase or decrease rate), he manifests changes in both speech and pause times (Gilbert, 1965; Lass, 1968; Minifie, 1963), there is the possibility that in his attempt to change the duration of his pauses, he will also manifest similar changes in his speech time as well.

However, it should be noted that the results of this investigation pertain to reading rate alone; caution must be exercised in generalizing results from reading to speaking situations. I am currently involved in the planning stages of a research project concerned with the effect of pause time alterations in impromptu speaking on perceptual judgments of speaking rate. Since there are some basic differences with regard to pause time between reading and speaking (Lass, 1968; Snidecor, 1943), such a study is necessary before statements concerning pause time alterations in speaking tasks and perceptual judgments of speaking rate can be made.

ACKNOWLEDGMENTS

This research was supported by Grant No. NB 05362-07 from the National Institute of Neurological Diseases and Blindness to the Bureau of Child Research of the University of Kansas. This research was conducted while the author was a post-doctoral Research Fellow of the Bureau. Special thanks are extended to Robert Mostellar, Department of Biometry, University of Kansas Medical Center, for his statistical consultation and advice, and to Ralph Shelton, Department of Hearing and Speech, University of Kansas Medical Center, for his helpful comments throughout the research.

CHAPTER XIII

EFFECTS OF TEMPORAL SPACING ON LISTENING COMPREHENSION: A SOURCE OF INDIVIDUAL DIFFERENCES

Raymond L. Johnson and Herbert L. Friedman*

Guilford's structure-of-intellect model posits the existence of 120 discrete intellectual abilities (Guilford, 1967). One of these, the ability to evaluate semantic relations (EMR) we have found to be a significant correlate of listening comprehension at high rates of compression (Friedman & Johnson, 1968, 1969). In this paper, results are reported which indicate that EMR is also related to a listener's ability to take advantage of a particular kind of perceptual aid in hearing compressed speech--the selective insertion of temporal spaces at major syntactic junctures within a sentence.

EMR as a Correlate of Listening Comprehension

According to Guilford's theory, the evaluative operation is a process of comparing and matching stimuli or items of information according to some criterion: identity, similarity, consistency, or conformity to rules for class membership. What is compared in the case of EMR are the implied relationships among two sets of words. One of the reference tests for this ability, Verbal Analogies, presents to Ss an item in the following form:

TRAFFIC: SIGNAL as RIVER: ______________ a. bank
b. dam
c. canal
d. sand bags

They are instructed to choose one of the four alternatives which is related to RIVER in the same way that SIGNAL is related to TRAFFIC. To answer correctly, the S must discover an attribute of both SIGNAL and TRAFFIC which is shared by RIVER and one of the four available choices. Specifically, the S must recognize that a signal stops the flow of traffic in somewhat the same way that a dam stops the flow of water. The EMR ability thus involves the recognition of implied relationships among items of information.

---

*Dr. Raymond L. Johnson is Research Scientist in the Communication Skills Research Program at the American Institutes for Research, 8555 Sixteenth Street, Silver Spring, Maryland 20910. Dr. Herbert L. Friedman is Director of the program.

Another characteristic which may define this evaluative ability is the rate at which the recognition occurs. It is historically significant that prior to Guilford's classification theory, the evaluative operation was variously termed perceptual speed, speed of judgment, and speed of association (Hoepfner, Nihira, & Guilford, 1966). A possible link between EMR and the comprehension of compressed speech suddenly becomes more plausible (and less suspect of being merely a statistical artifact) once we know that some type of perceptual or cognitive speed may be an important aspect of this ability.

Our finding that the magnitude of the correlation between EMR and listening comprehension is determined by the rate of compression suggests the hypothesis that understanding a highly compressed speech signal depends upon certain skills which are less clearly implicated at normal or near-normal rates. The possibility that some special competence is needed to comprehend highly compressed speech is consistent with findings from studies of individual differences in perceptual motor skills which have demonstrated changes, as a task becomes more difficult, in the relative contribution which specific skills make to the performance of the task (Fleishman, 1957).

The Effect of Temporal Spacing on Listening Comprehension

In a series of experiments we have demonstrated, as have others, that the comprehension of compressed speech is markedly improved by the selective insertion of temporal spaces at major phrase boundaries. One experiment used, as stimulus materials, sentences and sentence-like strings which were of three types:

> Meaningful, grammatical (for example): "Colorless cellophane packages crackle loudly."
> Meaningless, grammatical: "Colorless yellow ideas sleep furiously."
> Meaningless, ungrammatical: "Sleep roses dangerously young colorless."

Subjects listened to these strings either spaced at positions which conformed to the phrase structure, defined by an immediate constituent analysis:

> The clock / was built / by a Swiss watch maker,

or spaced to violate that structure:

> Union / leaders call sudden / strikes,

or not spaced at all. The strings were presented at normal rate and at three degrees of compression. The S's task was to repeat each sentence immediately after hearing it. We found, in analyzing the data, that while spacing did not improve recall of meaningless, ungrammatical strings--the random word sequences--spacing at phrase boundaries did improve recall of both meaningful grammatical and meaningless grammatical strings--especially when heard at the highest rate of compression, about 450 words per minute (wpm). For both of the grammatical string types there were no differences in recall between spacing which violated phrase structure and the absence of any spacing. Thus, temporal spacing helped the listener at high rates of compression--provided the string was well-formed grammatically and provided the spaces were inserted at phrase boundaries within the string. On the basis of these and similar findings we concluded that the function of spacing was to help the listener organize a sentence into easily remembered word groups which reflected the underlying syntactic structure of the sentence (Johnson, Friedman, & Stuart, 1970).

In subsequent studies we examined the effect of temporal spacing on the comprehension of connected prose passages, using cloze-type tests as measures of listening comprehension (Friedman & Johnson, 1969). To very briefly summarize the results of these studies, we found that listeners more accurately understood passages which were temporally spaced at phrase boundaries within each sentence than passages which were spaced between sentences (but not within), or passages not spaced at all. Furthermore, comprehension was better for sentences spaced at phrase boundaries than at clause boundaries. Phrase spacing produced consistently higher comprehension test scores than did clause spacing over various rates of compression. And clause spacing, in turn, produced higher scores than no spacing at all.

Purpose of the Study

Both lines of inquiry have led us into problems of interpretation. Evidence for the effects of temporal spacing has been somewhat troublesome to handle since there are at least two alternative explanations: the spaces may provide perceptual processing time for the listener, or may serve to mark syntactic boundaries for easier recognition. Likewise, the emergence of EMR measures as predictors of listening comprehension at high rates of compression is a finding of unappraised value in the absence of a usefully detailed general theory of listening comprehension with which this fact could be integrated. Since both the effects of temporal spacing and the predictiveness of EMR become discernible only at high rates of compression, and because EMR itself may involve a perceptual or cognitive speed component, we designed a study which included both as independent variables. The purpose was to determine whether the ability to evaluate semantic relations influenced the ability to use temporal spacing in understanding compressed speech. The joint manipulation of these variables was expected to help clarify one or both of their roles in listening comprehension.

Method

Subjects

Fifty-one undergraduates were recruited from local colleges and universities to serve as paid Ss: 25 were women, 26 were men. All were native speakers of English and were free from hearing or language deficits. Mean age was 22 1/2 years.

Passages

A magazine article about guitarist Charlie Byrd was divided into three approximately equal sections, designated passages A, B, and C. Three different versions of each passage were then prepared:

1. Temporal spacing between phrases. An immediate constituent analysis identified major phrase boundaries in each sentence, and taped versions of the three passages were recorded with temporal spaces inserted at every juncture. The narrator attempted to preserve natural intonation patterns in reading the sentences.
2. Temporal spacing between clauses. Sentences were so divided that clause units were kept intact. In recording this version of the three passages, temporal spaces were inserted to mark off major clause constructions without distorting the rhythm and intonation of the sentences.
3. No spacing. The three passages were recorded by the narrator who read at a "natural" pace without introducing prolonged pauses within sentences.

Each passage was compressed on the Tempo-Regulator to a rate 2.0 and 2.75 times normal.

Comprehension Measure

A cloze test was constructed for each passage by randomly deleting 20 lexical words. The format of the test was a typed transcript of the passage with blank spaces where deletions occurred.

Ability Measure

As a measure of EMR, the ability to evaluate semantic relations, every S received the Verbal Analogies test from the Guilford battery of experimental measures. Previous multiple correlation studies had found that the Verbal Analogies test emerged as a significant predictor of listening comprehension at high rates of compression (Friedman & Johnson, 1969).

Procedure

Subjects were randomly assigned to nine experimental groups of approximately equal size. Groups I, II, and III heard the three passages without any temporal spacing within sentences. Groups IV, V, and VI heard the passages with spacing at phrase boundaries, while Groups VII, VIII, and IX heard the passages with spacing at clause junctures. Every S heard one passage at each of the three rates, and the order of presentation was always the same: first normal, then 2.0 times normal, and finally 2.75 times normal. For each spacing treatment, the passages were counterbalanced so that every passage was heard at every rate. The experiment thus employed 27 different tapes (three passages x three spacing conditions x three rates of presentations). The design is summarized in Table 13.1. All Ss in an experimental group participated at the same time. A tape was played to the group and immediately at its conclusion the comprehension test based on that passage was administered. Three tapes and tests were given in a single session.

TABLE 13.1

DESIGN OF THE STUDY. PASSAGES ARE DESIGNATED A, B, AND C

		Normal Rate	2.0 x Normal	2.75 x Normal
No Spacing	G_1	A	B	C
	G_2	B	C	A
	G_3	C	A	B
Phrase Spacing	G_4	A	B	C
	G_5	B	C	A
	G_6	C	A	B
Clause Spacing	G_7	A	B	C
	G_8	B	C	A
	G_9	C	A	B

Results

Analyses of variance were performed to determine the effects of EMR and presentation rate on the comprehension of passages (a) temporally spaced at phrase boundaries, (b) temporally spaced at clause boundaries, or (c) not spaced within sentences.

Statistically significant findings are enumerated briefly below:
1. Both EMR and presentation rate were significant main effects under all spacing conditions. Compressed passages were never understood as well as passages presented at normal rate, and Ss above average in EMR ability (i.e., those whose scores fell above the group mean) always did better on comprehension tests than below average Ss.
2. Temporal spaces inserted at phrase boundaries always resulted in more accurate listening comprehension than did clause spacing (for both normal and compressed rates of presentation and both levels of EMR ability). Spaces introduced at clause boundaries always resulted in better comprehension than not introducing any spaces within sentences.
3. The smallest difference between high and low ability listeners was found when passages were phrase-spaced and presented at normal speed.
4. The greatest difference between high and low ability listeners was found when passages were phrase-spaced and presented at a rate 2.75 times normal.
5. Whatever the nature of the help phrase spacing gave to listeners, those below average in EMR were most benefited when the spaced passage was presented at normal rate, while those listeners who were above average were helped most in coping with highly compressed spaced passages.

These data are graphically presented, in a somewhat simplified form, (i.e., the twice normal rate has been omitted) in Figure 13.1.

Discussion

The insertion of temporal spaces at phrase boundaries enabled listeners of below-average EMR ability to understand passages at normal speed almost as well as above-average listeners. This finding implies that the provision of spaces in some way compensated--at normal speed--for the limitations in the ability to evaluate semantic relations. Comprehension for listeners above average in EMR ability, on the other hand, was little influenced by temporal spacing at normal speed, but was significantly enhanced at high rates of compression. For these listeners, the insertion of temporal spaces compensated, in part, for the loss of perceptual processing time or the loss in intelligibility due to speech compression. One conclusion to be drawn from these findings is that the level of EMR ability enters into determining the amount of time a listener needs in understanding a message. While high ability listeners need additional

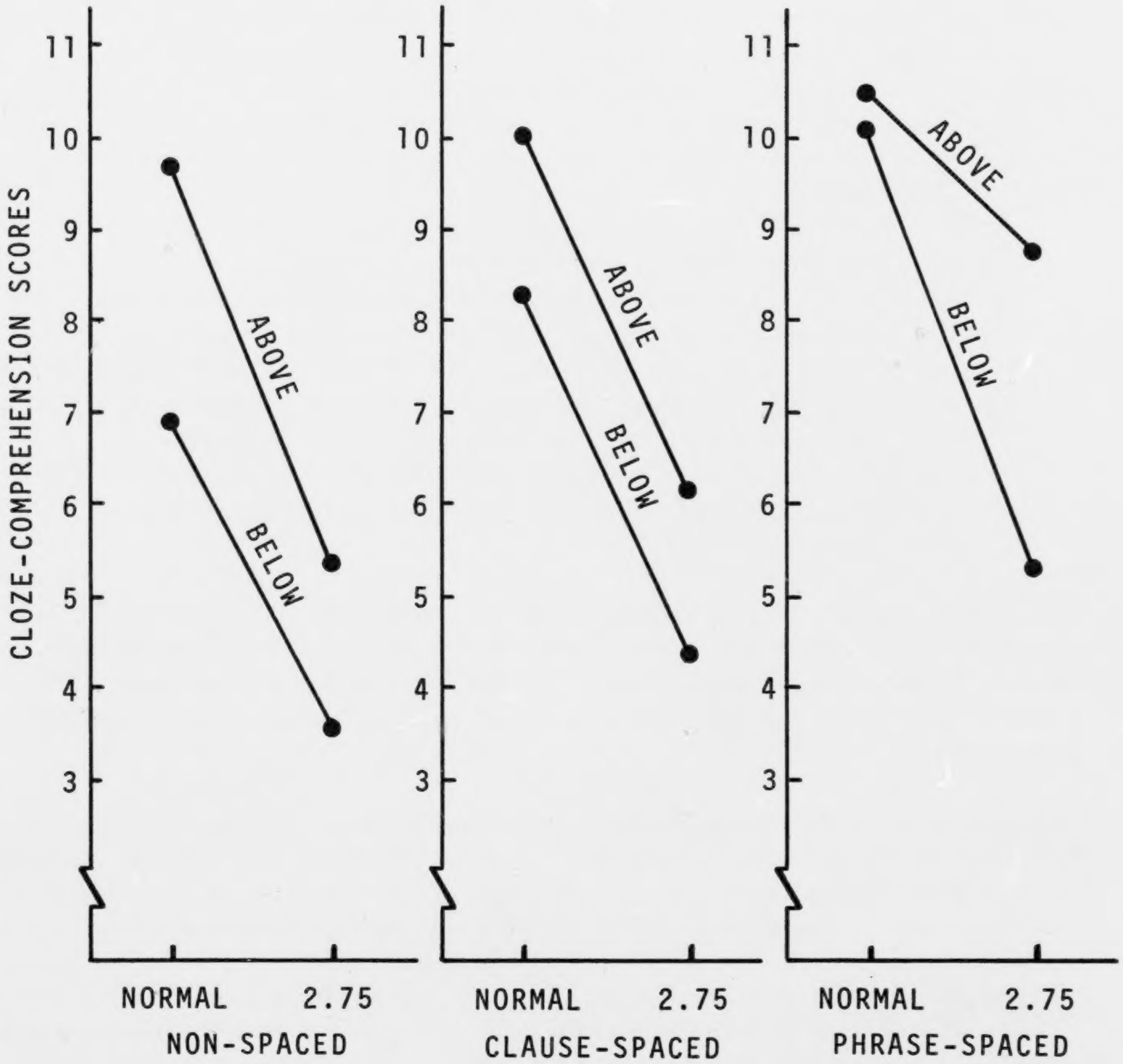


Figure 13.1. Graphical comparisons of cloze comprehension scores for subjects scoring above and below the group mean on the Guilford Verbal Analogies test. Note that data for passages presented at twice normal rate have been omitted in order to condense the graph sizes.

time at phrase boundaries only for highly compressed speech, listeners of low ability can use the added time even at normal rate.

Measures of EMR ability, such as Verbal Analogies, may be sensitive indicators of a listener's time needs because the solving of the test problems involves the same mental operation required for comprehension. Since proficiency and speed usually go together, people who are adept at evaluating semantic relations do so quickly. Those less adept find solutions more slowly. Listening comprehension, as EMR, involves the recognition of relationships: relating sounds with sounds, words with words, phrases with phrases, sentences with sentences. As Lenneberg (1969) has noted:

> ". . . virtually every aspect of language is the expression of relations. This is true of phonology (as stressed by Roman Jakobson and his school), semantics, and syntax. For instance, in all languages of the world words label a set of relational principles instead of being labels of specific objects. . . . Further, no language has ever been described that does not have a second order of relational principles, namely, principles in which relations are being related, that is, syntax in which relations between words are being specified [pp. 640-641]."

There are relationships among sentences also, as attested by the fact that readers can identify paragraph boundaries in unindented prose passages (Koen, Becker, & Young, 1969). Comprehension as a process can be understood, then, as the recognition of relationships, within and between sentences.

From our everyday experiences as listeners we can observe that the difficulties we have in understanding what someone has said is not usually in comprehending sentences taken separately, but rather in seeing how one sentence relates to another. Comprehension falters when a listener cannot recognize the way the speaker couples sentences to form a train of thought. Some evidence in support of this view is provided by Brent (1969) who tested an hypothesis that the integration of sentences into paragraphs is a process which unfolds in time. His findings indicate that a person may completely comprehend the meanings of each of the sentences in a paragraph, as coherent functional units, without being able to "grasp" or "utilize" the integrative structure of the paragraph as a whole. This act of integration requires time between sentences (perhaps of the magnitude of 1 to 3 seconds, Brent has roughly estimated), an experimental finding which further strengthens our contention that the ability to evaluate semantic relations may determine the amount of time a listener needs to understand a message. It is reasonable to agree further, on the basis of evidence, that a more profound view of "readability" or "listenability," and the psycholinguistic factors which influence these, would involve the nature of the relationships among sentences in a text, at least as much as variables such as word frequency and length, or sentence length and complexity.

The recognition of relationships at the phrase and sentence level has been incorporated as a central component in a theory of language comprehension developed by Quillian (1966, 1967, 1969). As Quillian depicts the process:

> "It seems generally agreed that understanding text includes recognizing the structure of relations between words of the text. . . . The overall effect of these processes is to encode the text's meaning into some form more or less parallel to that in which the subject's general knowledge is stored, so that its meaning may be compared to that knowledge, and perhaps added to it [Quillian, 1966, p. 53]."

Quillian's "path finding" theory of comprehension is too detailed for us to attempt more than a cursory exposition here. In broad outline, however, the theory states that words are encoded in semantic memory as attribute-bundles and these attributes are extensively cross-indexed. The interlocking network of attributes which characterizes the organization of semantic memory permits two words to be compared for shared meaning by searching through their respective attribute fields to discover any intersect or overlap in attributes. Sometimes two words will not share any immediate attributes, but both will be related--in different ways--to some third word. Quillian suggests that the length of the search required to find a path connecting two words is an index of their semantic similarity. If the path is long and circuitous, a sentence in which these words co-occur may be difficult to understand. If a path cannot be found which links the words of a sentence, the sentence is meaningless. In Quillian's view:

> "The cognitive processing which a reader must carry out in order to [understand a text] is based on his finding, for certain pairs . . . of concepts which the text associates, some way in which those same concepts previously have been, or intelligibly may be, associated, given his general memory [Quillian, 1966, p. 70]."

Path finding makes use of the syntactic information in a sentence in several ways:

1. To form units. Data from Quillian's simulation studies suggest that readers or listeners "bite off" segments of text for intensive processing. Left-to-right segmentation is achieved by applying unit-forming rules which are purely syntactic in nature, and results in phrase groupings which contain two or three lexical items. Our own finding that the insertion of temporal spaces at phrase boundaries facilitates comprehension is compatible with this aspect of Quillian's theory.

2. To determine the order in which words in a unit are related. Once a phrase has been isolated for processing, syntactic information is used in deciding the order in which the attribute fields of constituent words are to be searched.
3. To relate one phrase or sentence to another. Following the processing of all words in a phrase or sentence, it is linked to the preceding portion of the text, and this higher-order connection is guided by syntactic cues.

On the basis of these considerations, we may formulate the following hypothesis: that EMR, the ability to evaluate semantic relations, is involved in the pathfinding process essential to the comprehension of connected discourse, and that the Verbal Analogies test measures individual speed and proficiency in finding connecting paths between concepts.

ACKNOWLEDGMENT

The research reported herein was supported by the United States Office of Education under Grant No. OEG-1-7-070642-3764.

CHAPTER XIV

THE COMPREHENSION OF RATE INCREMENTED AURAL CODING

Murray S. Miron and Eric Brown*

An earlier publication (Miron & Brown, 1968) has reported the preparation of a battery of stimulus tapes in which three parameters potentially critical in their control of comprehension were conjointly manipulated. These parameters were: (1) talker rate, manipulated through the limits attainable by a trained speaker; (2) selective pause compression in which pause time was proportionally deleted; and (3) random message deletions. The results of that earlier study had indicated that comprehension would not necessarily be additively affected by these differing means of achieving an increase in message rate. For the first time, message rates approaching 1,000 words per minute (wpm) were achieved with only relatively moderate sacrifice of total phonation time. The earlier article had as its purpose the reporting of the methods by which these rates could be achieved. The present report presents comprehension data and psychophysical judgments on the extremely broad range of speech rates represented by these stimulus materials.

Earlier research (Fairbanks, Guttman, & Miron, 1957c) confirmed by a number of other investigators (see Foulke & Sticht, 1969 for a complete review of the literature) had indicated that beyond compressed rates of approximately 275 wpm there is a precipitous decline in comprehension. Rates of approximately 160 wpm have been found to be normal and preferred (Hutton, 1954). Thus, as yet no appreciable gain in listening efficiency without sacrifice of comprehension has been achieved by the speech compression technique. Even assuming that with cleaner signal-to-noise characteristics and a moderate tolerance of loss of comprehension, doubling the normal speech rate by means of random deletions still would not produce an absolutely large increase in listening efficiency. What is required is an entirely different approach to the problem; namely a combination of selective and random speech deletions.

---

*Dr. Murray S. Miron is Associate Professor of Psycholinguistics, Psychology Department, Syracuse University, Syracuse, New York 13210. Eric Brown is research assistant to Dr. Miron.

Much of the normal speech signal is redundantly specified. Some 30% of the normal reading of an extended prose text can be identified as pause time (Miron & Brown, 1968). Approximately 50% of total message time can be discarded without appreciable loss in comprehension. And a speaker can increase his normal rate of talking by at least 30% without listener loss of comprehension. The problem is in identifying what precise parts of the speech signal may be discarded. At first thought, it seems reasonable to suppose that since pause time merely adds to message length, pause deletion should increase listening efficiency without loss of comprehension. But the speech pause does contribute to message comprehension. In spontaneous speech, the pause marks retrace corrections, signals ideational boundaries, and high information points (Goldman-Eisler, 1968; Maclay & Osgood, 1959). In reading, its role is less well specified, but it is reasonably obvious that pauses are at least grossly correlated with sentential and ideational syntax. The good reader does not pause because he is out of breath; the text dictates where he must breathe. This report provides a preliminary exploration of the roles of pause time, talker rate, and total message duration in comprehension, and their effects on the judgment of rate. Such data are preliminary to a fuller understanding of the factors which contribute to the comprehension of speech and to the practical problem of achieving significant increases in listening efficiency.

Method

The 36 stimulus tapes as prepared by the methods already described in an earlier publication (Miron & Brown, 1968) were each presented to 36 haphazardly assigned groups of five Ss (45 to 60 Ss per treatment). Subjects were all native speakers of English drawn from the introductory psychology course of Syracuse University. Subjects listened to the tapes through Sharpe HA-8A earphones driven by a Magnecord 1022X tape recorder with a received signal at approximately 60 db sensation level. An instruction tape preceded each experimental session in which the history and purpose of speech compression were discussed and examples of actual compression were given. Immediately following the message, a 55-item comprehension test was administered. All stimulus tapes represented various treatment manipulations of a technical message on meterology for pilots. The overall rates of these 36 conditions varied from 129.1 to 854.8 wpm. These rates were achieved through the combination of: (1) original talker rates (T_R) of 129.1, 164.2, and 224.0 wpm, representing the slowest (TR/S), normal (TR/N), and fastest (TR/F) rates the trained speaker could sustain without sacrifice of articulatory accuracy or effective delivery style; (2) excision (P_c) of 100%, 50%, or 0% of all pause lengths in excess of 50 milliseconds (msec.); and (3) periodic deletion (R_c) of approximately 40 msec. invervals of the total message duration at intervals appropriate for the reduction of the total message time by 0%, 30%, 50%, or 70%.

An additional group of 18 *S*s was presented with a 150-word excerpt of the message at each of the 36 condition speeds and asked to estimate on a 7-point scale:

FAST _____:_____:_____:_____:_____:_____:_____ SLOW

the relative speeds of these conditions. Anchor points representing the slowest and fastest rate conditions were provided in the instructions as examples.

Results and Discussion

Treatment Effects upon Comprehension

Table 14.1 details the variance contributions of each of the experimental variables to listening comprehension. On a dichotomous decision basis (significant/nonsignificant), all three methods of increasing rate significantly ($p < .01$) decrease comprehension. The interaction of talker rate and random compression also produces significant nonparallel changes in comprehension ($p < .01$); otherwise the combination of the three variables only linearly displaces comprehension scores. On the other hand, if one uses the estimates of variance contributions to the control of comprehension (omega-square), it is strikingly clear that random compression controls almost all of the reliable variance (64%). Obviously, the particular choice of values of the independent variables will effect the contribution of those variables. But the levels of the variables chosen for talker rate, pause compression, and random compression were not arbitrary in this study. Random deletion of 70% of a message has been found to be about the practical limit to the speech compression process. Beyond this point, speech fragmentation and inherent system noise become damagingly obtrusive and comprehension scores drop to or below chance levels (Fairbanks et al., 1957c). Sustained talker rates in excess of the TR/F condition of this experiment (224 wpm) are difficult to produce without sacrifice of articulatory precision. Short bursts of speech in excess of this value can be produced, but they require considerable editing and re-recording in longer messages. (With considerable practice, for example, Fairbanks was able to produce a rate of 344 wpm for a 55-word passage, see Hutton, 1954). But these higher talker rates effect prosodic, intensity, and fundamental frequency attributes of the reading in increasingly abnormal ways. The pause excision procedures in the limiting case of the PC/100% condition removed all pauses of 50 msec. effective duration. Thus, all three variables have been manipulated through their approximate effective ranges. When the combination of random and pause compression effects are held constant, a talker rate increase of 181 wpm (.21 log unit) is possible. A 71 wpm increase in overall rate (.08 log unit) can be achieved by pause compression and a 445 wpm increase (.52 log unit) by random compression when the remaining combinations of effects are held constant.

Simultaneous application of the limits of all three variables effects an overall increase in rate of 734 wpm (.83 log unit). These effective word rate increments should be compared to the approximate range of 250 wpm employed by most other investigators (see, for example, Fairbanks et al., 1957c; Foulke & Sticht, 1967a; Orr & Friedman, 1964). In sum, although the decrement in comprehension produced by talker and pause compressions is significant, the loss is moderate when compared to the obtained increment in listening rate.

TABLE 14.1

VARIANCE CONTRIBUTIONS OF TREATMENT EFFECTS ANALYSIS OF VARIANCE TABLE

Source	df	SS	MS	F	ω^2
I Effect	2	0.2253	0.1126	8.4415	.03
J Effect	2	0.1349	0.0674	5.0540	.02
K Effect	3	5.0850	1.6950	127.0320	.64
I x J	4	0.0277	0.0069	0.5190	0
I x K	6	0.3826	0.0638	4.7784	.05
J x K	6	0.0351	0.0058	0.4382	0
I x J x K	12	0.1148	0.0096	0.7170	.01
Within	144	1.9214	0.0133		
Total	179	7.9267			

Key:
I = talker rate
J = pause compression
K = random compression

Treatment effect means are displayed in Table 14.2.

Table 14.3 displays the least squares equations relating the log word rate and comprehension probability. The values superimposed on this table represent, respectively: the mean value of the x-variate, (XB); the mean of the y-variate (B); the standard deviations of x and y, (SX, SY); the Pearson product-moment correlation between x and y, (R); and the best linear least square fit of y as a function of x, (YP = aX+b).

It will be observed that log word rate represents an exceedingly good fit to comprehension probability (r = .90). It should also be observed that the comprehension instrument has a mean value over all conditions which is quite close to the centered value of 60% between chance performance (20%) and maximum (100%). As a consequence, ceiling or floor effects in the dependent variable should not be affecting the obtained relationship.

TABLE 14.2

SUMMARY OF EXPERIMENTAL VARIABLES FOR TREATMENTS

Condition	Rate (wpm)	Rate (wpm)	Judged Rate	Predicted Rate	Comp. (%)	Comp. (Z)
TR/ SLOW	287.52	2.46	4.63	4.41	0.55	0.15
TR/ NORM	352.90	2.55	5.02	5.16	0.54	0.11
TR/ FAST	468.78	2.67	5.73	6.18	0.47	-0.27
PC/ 0	335.67	2.53	5.04	4.90	0.53	0.06
PC/ 50	366.98	2.56	5.01	5.23	0.55	0.15
PC/ 100	406.55	2.61	5.33	5.62	0.48	-0.21
RC/ 0	190.29	2.28	2.83	3.22	0.69	0.95
RC/ 30	272.49	2.44	4.69	4.51	0.61	0.47
RC/ 50	380.82	2.58	6.04	5.72	0.53	0.05
RC/ 70	635.33	2.80	6.94	7.56	0.25	-1.47

TABLE 14.3

LINEAR TRANSFORMS OF COMPREHENSION AS A FUNCTION OF LOG WORD RATE

	Log Rate (WPM)	Scaled Comp.					
	$\overline{X}$	$\overline{Y}$	S_X	S_Y	$Y_P =$ a	(x) + b	r
TR/ S	2.41	.69	.21	.24	-1.02	3.16	-.89
TR/ N	2.50	.68	.21	.25	-1.11	3.46	-.93
TR/ F	2.63	.59	.21	.23	-1.04	3.32	-.94
PC/ 0	2.47	.67	.23	.24	- .94	2.99	-.88
PC/ 50	2.51	.69	.22	.24	-1.04	3.31	-.94
PC/ 100	2.56	.61	.22	.23	- .92	2.96	-.88
RC/ 0	2.27	.88	.10	.09	- .58	2.19	-.64
RC/ 30	2.42	.77	.10	.07	.04	.68	-.05
RC/ 50	2.57	.67	.10	.15	-1.37	4.19	-.91
RC/ 70	2.80	.31	.10	.04	- .16	.77	-.44
Overall	2.52	.66	.22	.19	.96	3.08	-.90

(Detailed factor and item analyses of the comprehension instrument, although performed, are not included in this report; the critical point is that the instrument is both reasonably reliable and discriminating). That comprehension should be found to be a lawful function of log word rate is not surprising. But it is important to determine whether changes produced by the three methods are individually fitted by this general function. The overall relationship between comprehension and word rate is closely approximated when pause and talker subconditions are held constant. When talker and pause compression effects are allowed to vary while random compression is held constant (by taking variable means over those conditions sharing a specified R_C), rate changes are not a good linear fit to comprehension. A closer inspection of the data (see Table 14.2) reveals the reason for the discrepancies. The slowest talker rate does not produce the highest comprehension. In fact, the highest comprehension is obtained when the slow talker rate is treated by pause reduction of 50%. Even a 100% pause reduction of the normal rate message has a comprehension score which is within one-hundredth of a percent of the score obtained for the untreated normal rate message. Pause compression actually improves a slow rate message and can be used to increase rate without the to-be-expected loss in comprehension predicted by the increment in that rate. There are, of course, limits to this process. The rate effects upon comprehension reemerge when the compression is held constant at 50%. But the worst fit of the overall relationship between comprehension and rate is obtained when RC is held constant at the moderate value of 30%. At this level, rate increase are achieved by talker and pause treatments without appreciable loss in comprehension. By the time R_C is increased to 70%, comprehension begins to approach its chance asymptote and the facilitative effects of the talker and pause treatments are lost in the noise.

Rate Judgments

The disparate effects of talker and pause treatments can also be observed in the judgment of message rate. Table 14.4 displays the relationships between log word rate and judged message rate for all treatment combinations. Again, it is not surprising to find that, overall, judged rate is a lawful function of log rate in wpm. When the predicted rate as determined from the least squares fit of judged rate to log measured rate (dropping the truncated points) is related to comprehension, the prediction of comprehension is at least as accurate as that for log word rate. The lawful relationship between judged rate and comprehension has the additional advantage, however, of indicating that the stimulus measurement of rate in wpm has psychological meaning through the mediation of its linear relationship to judged rate.

TABLE 14.4

LINEAR TRANSFORMS OF JUDGED RATE AS A FUNCTION OF LOG WORD RATE

	Log Rate (WPM) $\overline{X}$	Judged Rate $\overline{Y}$	S_X	S_Y	$Y_P =$ a (x)	- b	r
TR/S	2.41	4.63	.21	2.09	9.51	18.34	.95
TR/N	2.50	5.02	.21	1.68	7.70	14.27	.94
TR/F	2.63	5.73	.21	1.25	5.50	8.73	.91
PC/0	2.47	5.04	.23	1.81	7.37	13.18	.92
PC/50	2.51	5.01	.22	1.83	7.62	14.14	.92
PC/100	2.56	5.33	.22	1.64	6.92	12.38	.92
RC/0	2.27	2.83	.10	.92	8.26	15.91	.89
RC/30	2.42	4.69	.10	1.09	10.39	20.50	.96
RC/50	2.57	6.04	.10	.35	.66	4.34	.19
RC/70	2.80	6.94	.10	.06	.47	5.64	.76
Overall	2.52	5.13	.22	1.72	7.20	12.98	.92

As was the case with the treatment effects on comprehension, judgments of talker and pause treatment rates are not well estimated by the overall function relating judged rate to measured rate when compression treatments are held constant. In this instance, there is least change in the estimations of rate, at a constant random compression of 50%, with talker and pause-controlled increments in actual rate. Assuming a linear relationship between log rate changes and judged rate changes (the obtained relationship was .92), increments in actual rate by means of talker or pause treatments simply are not detected by the observer when moderate random compression treatments are employed.

Message Effectiveness and Efficiency

The best summary which can be given to the data so far presented is to relate the treatment effects to the efficiency and effectiveness of the listening conditions. Message effectiveness is defined as the obtained comprehension score less chance performance level, scaled to range from 0 to 100%. That is, the effect of the message on comprehension scores above that which could be achieved at 100% compression. Message efficiency is defined as the ratio of message effectiveness to message duration. For the purposes of the following display, efficiency is scaled to represent percent comprehension relative to the maximum efficiency for this message.

Figure 14.1 displays the relationship between message effectiveness and message efficiency graphically. It will be observed that the effectiveness index has the expected ogival shape. Message content is sampled as a random proportion of the content as yet unlearned. Thus, at slow rates, complete comprehension of the message recapitulates Xeno's paradox. The limiting value of 100% comprehension is reached by proportions of the as yet remaining content. The expected value of the distributions of proportionate content samples varies as a function of the log message rate; on the average, smaller proportions of the content are sampled as the rate is increased, larger proportions as the rate is slowed. That is to say, the distribution of comprehension is a log-normal function of rate. The implication of this assertion is far-reaching. It implies that the effectiveness of a message is relatively fixed; only the efficiency with which such an effectiveness is achieved can be increased. Fairbanks et al. (1957c) found this to be the case. In an attempt to utilize the message time saved by a 50% compression, the investigators selectively augmented the message content so that the total message time equalled the uncompressed duration; the augmentation simply repeating certain of the passages already contained in the message. They found no increment in comprehension over that for the unaugmented message. Random proportions of the message content were still being sampled.

Turning to the analysis of message efficiency, it will be observed from Figure 14.1 that efficiency rises to a peak at approximately 2.6 log units of rate (400 wpm). This peak of 400 wpm is to be compared with a peak of 280 wpm as found by most other investigators (see particularly Fairbanks et al., 1957c). The striking difference in the efficiency maxima is due to the difference in the methods of achieving increased rates. The message content is precisely the same as that used by Fairbanks et al. (see Miron & Brown, 1968). This experiment, however, employed selective compression in addition to the random compression which had previously been used. Aural coding rates which begin to approximate reading rates are, by these techniques, closer to achievement.

ACKNOWLEDGMENT

This research was supported by Grant RO1-HD02417 from the National Institute of Child Health and Human Development.

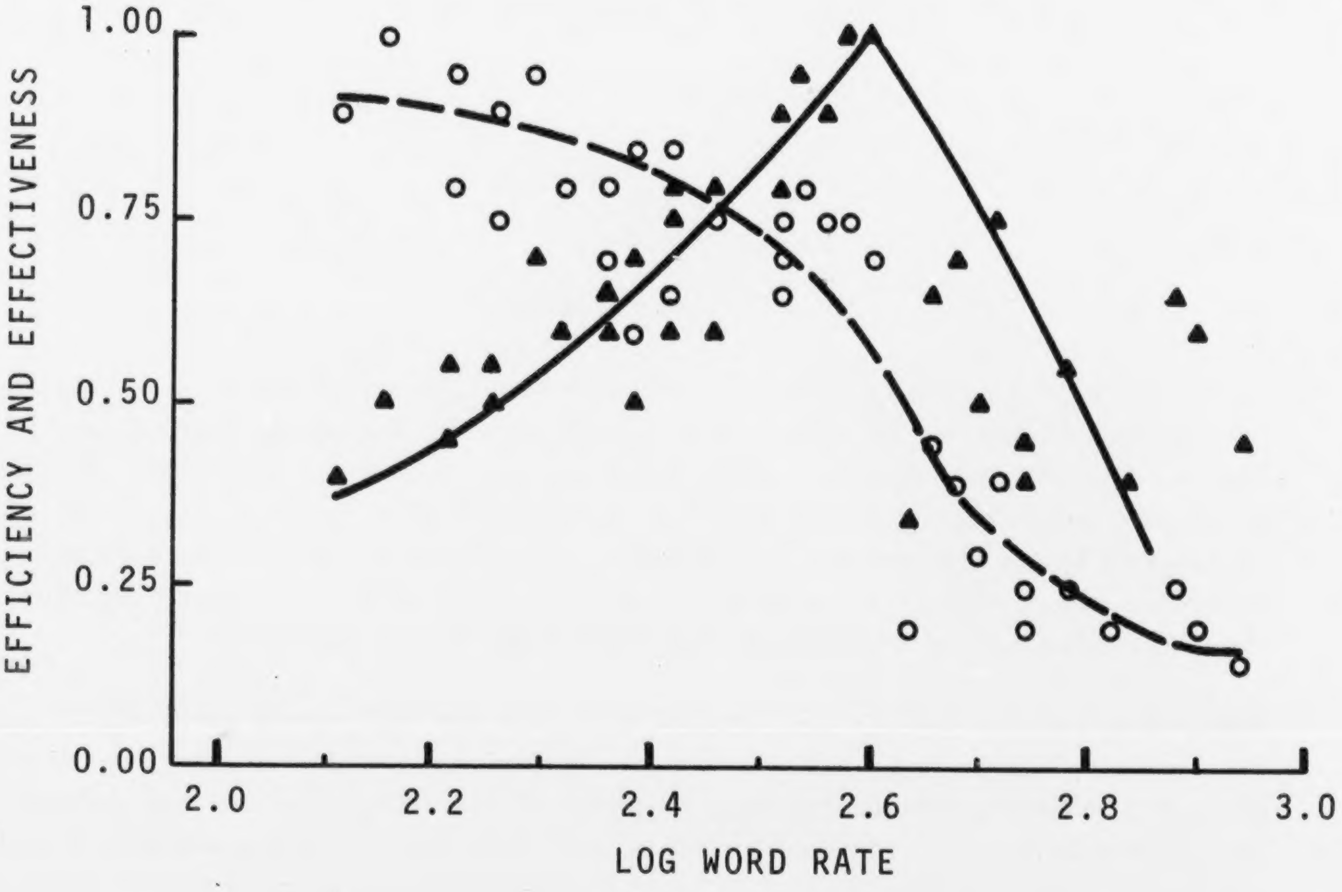


Figure 14.1. Message efficiency and effectiveness as a function of log word rate.

CHAPTER XV

THE EFFECTS OF SOURCE LANGUAGE PRESENTATION RATE ON THE PERFORMANCE OF SIMULTANEOUS CONFERENCE INTERPRETERS

David Gerver*

Introduction

In simultaneous interpretation, as in most naturally occurring tracking tasks, the observer is often confronted with differential information load. For the interpreter, this may be due either to syntactic and/or semantic variability of the source language input, and/or to variability in source language presentation rate. In the present study, attention will be paid to the effects of presentation rate of the source language on the performance of simultaneous conference interpreters.

Goldman-Eisler (1968) and others have shown that most periods of speech consist not only of speech but also of silent intervals of varying temporal duration. Goldman-Eisler has suggested that the more of his output that the simultaneous interpreter can crowd into the source speaker's pauses, the more time he has to listen to the input without interference from his own output. Unfortunately for the interpreter, though, it is doubtful whether he can reliably predict input pauses and cannot achieve the "ideal" distribution of his own speech time and pause times, i.e., to pause when there is input speech and to speak when there are input pauses. Even if he could do this, it is doubtful whether the simultaneous interpreter could cram much of his own output into input pauses, since the majority of pauses in speech are less than 0.5 second in duration, while only 20% to 40% lie between .5 and 1 second, 12% to 20% between 1 and 2 seconds, and very few above 2 seconds (Goldman-Eisler, 1961c).

While it is feasible that the interpreter could utilise input pauses, in view of the above mentioned limitations it seems more likely that he would attempt to cope with increase in input rate in other ways. He might, for instance, speak more and pause less, and increase his own output rate as input rate increased. Alternatively, he could pause more frequently and speak for shorter intervals at faster rates.

---

*Dr. David Gerver is a Lecturer in Psychology in the Department of Psychology at the University of Durham, Durham, England.

The purpose of this experiment, then, is to examine the effect of variation in input rate on the interpreter's performance by systematically varying the rate of presentation of a source passage. In order to ascertain the extent to which any variability in interpreter's performance may be due to difficulty in simply transmitting speech at faster rates rather than to difficulty in carrying out the complex decoding and encoding processes involved in interpretation as input rate increases, all relevant measures of the interpreter's performance will be compared with those for Ss shadowing (repeating as they hear it) the same experimental message. Carey (1968) found that shadowers made more errors as input rate increased from 1 to 3 words per second (wps). Treisman (1965), using statistical approximations to French and English and presentation rates of 1. 7 and 2. 5 wps, found that information rate had a greater effect on the number of correct responses produced by simultaneous interpreters than by shadowers. Treisman also found that ear-voice span (the number of words S follows behind the speaker) was greater for interpreting than for shadowing. She attributed the differences in performance between interpreters and shadowers to the " . . . increased decision load between input and output required in translation: two selections need to be made, the first to identify the word or phrase heard, and the second to select an appropriate response. The shadowing task is simpler if it is assured, as is plausible, that a single central identification of the verbal unit serves for both reception and response, so that only one decision is required. "

Treisman did not attempt to analyse her Ss' errors, but Carey employed four error categories: word omissions, word substitutions, additions of words, and distortions of words. Preliminary analysis of the protocols of Ss in the present study suggested that the term "discontinuity" rather than "error" was a more appropriate description of deviations from the input message found in the output of both interpreters and shadowers. Though omissions, repetitions, additions, and distortions can be regarded as errors, other phenomena found both in shadowing and interpretation, such as substitutions or corrections of words or phrases, are not necessarily errors but do involve some discontinuity in the message being transmitted. In the present study, the following categories will be employed: omissions of words, omissions of phrases, omissions of longer stretches of input of eight words or more, substitutions of words, substitutions of phrases, corrections of words, and corrections of phrases.

Substitutions involve approximate or less precise responses which, though grammatical and meaningful, alter the meaning of a sentence in some way. Corrections are observed whenever S interrupts his output to correct previous words or phrases and are of particular interest in simultaneous interpretation for the light they shed on feedback mechanisms. Welford (1968) discusses the simultaneous interpreter's performance within the context of a "single channel" hypothesis and the role of feedback in human information processing during continuous tracking tasks. He states that simultaneous interpreters can, after long practice, acquire the ability to speak and listen concurrently because they learn to ignore the feedback from their own voices. The very fact of frequent corrections in the interpreter's output, however, shows that though interpreters may learn to ignore the sound of their own voices, they do not ignore the meaning of what they say.

Apart from these "discontinuities" in output, the dependant variables in this study will be: the number of words correctly shadowed or translated, ear-voice span, utterance times, and unfilled pause times. Fries (1952) defined an "utterance unit" as " . . . any stretch of speech by one person before which there is a silence on his part, and after which there is also silence on his part." For the purposes of this study, an utterance will be defined as any period of speech bounded by unfilled pauses, the criterion for an unfilled pause being a break in the speaker's utterance of not less than 250 milliseconds (msec.). Goldman-Eisler (1968) adopted this criterion, arguing that pauses up to 250 msec. might occur as part of ritardando effects or articulatory shifts between plosives.

Finally, the ratio of overall pause time to overall speech time will also be calculated since it has been hypothesized that the interpreter might try to redistribute his performance in time as input rate increases. Apart from redistributing speech time and pause time, the interpreter might become less variable in his output rates as information load increases. One noticeable characteristic of the delivery of conference interpreters is the frequency of ritardando and accelerando passages. That is to say that the interpreter will appear to dawdle over some words, rather like a person thinking of something else while he is speaking, and then speak very quickly as if under pressure to unload material in store. This, in turn, may be followed by a further ritardando passage, and so on. Another way of describing this type of output is in terms of the variability of the relationship between the number of words uttered and the time taken to utter them. In ritardando passages the interpreter may utter a few words in a given time, whilst in accelerando passages he would utter more words in the same time. At slower input rates, then, the interpreter may have time to vary his output rate, but not at the faster rates, and this would be reflected in the correlations between the number of words per utterance and the time per utterance at different input rates.

Method

A French text of 550 words (an extract from a speech at a United Nations Educational, Scientific, and Cultural Organization Conference on Human Rights) was recorded on tape at a rate of approximately 120 words per minute (wpm) by a male native French speaker. This master tape was systematically expanded and compressed in time by means of an Eltro Tempophon, the rate being changed at intervals of approximately 110 words. The output of the Tempophon was then rerecorded on a Revox G36 tape recorder, and the final experimental tape contained a continuous text with the following rates and passage lengths:

Rate wpm	95	112	120	142	164
Number of words	108	106	108	111	118

Subjects

The Ss were 10 professional simultaneous conference interpreters. Five of the Ss were allotted to the shadowing condition, five to the interpreting condition. The mother tongue of all Ss was English.

Procedure

All Ss heard the stimulus tape under language laboratory conditions. The experimental tape was relayed to Ss' individual booths from the main control booth and was recorded on the top track of each S's tape. Subjects' responses were recorded on the bottom tracks of their tapes.

Interpreter Ss received the following prerecorded instructions: "You are going to hear a speech in French. You will probably notice that the speaker speaks more quickly as his speech continues. Please interpret the passage from French to English as you hear it."

Subjects in the shadowing group received the same instructions except that they were asked to repeat the words in French as soon as they heard them.

Treatment of Results

A 2 Groups x 5 Presentation Rate repeated measurements design (Winer, 1962) was employed in analysis of variance of the results.

In order to measure Ss' utterances and pause times, both tracks of each S's tape were transcribed on a paper record of a pen tracing of each channel. The pen recordings were obtained by replaying each S's tape on a Revox G36 stereo tape recorder, and feeding the output from each channel to two speech trigger units attached to a modified Marconi Electroencephalograph (EEG) pen recorder. In order to minimise pen onset and offset delays, which vary directly with signal level (Ramsay & Law, 1966), the output from each channel of the trigger circuit was monitored on a multichannel display oscilloscope, whilst the tape recorder output was monitored on separate loudspeakers. The auditory signals could then be matched with visual traces of the operation of the trigger units to produce optimal onset sensitivity and minimal offset delay, by adjusting the sensitivity of the trigger circuits. This circuit is illustrated in Figure 15.1.

After setting the appropriate level controls, the E followed the text from a typewritten copy and activated a marker pen on the EEG recorder at approximately every fifth word of the input text. As can be seen from Figure 15.2, this practice provides a number of reference points against which to match the tape recording with the pen recording of each channel. The speed of the pen recorder transport was 3 centimeters per second. Both tape-recorded tracks were transcribed onto the pen recordings, and the measurements of ear-voice span, speech time, and pause time were then made.

Counting the number of words correctly shadowed was a straightforward task; but in assessing the correctness of interpretations, paraphrase was taken into account since a word-for-word translation was not expected and, indeed, would not have been a good translation from the interpreter's point of view.

Ear-voice span was calculated for shadowers at every fifth word of the input text in terms of the number of words not yet correctly repeated by the S. Words omitted entirely in shadowing were counted as part of the ear-voice span until the S had passed beyond the point at which they could have been repeated in the correct order. For interpreters, ear-voice span was calculated at every fifth word of the input text in terms of the number of words not yet translated by the S. Words omitted in translation were counted as part of ear-voice span until the interpreter had passed beyond the point where they could have been translated in context. Provided that some reasonable connection could be inferred between the interpreter's output and the original message, an error in translation was counted as a word translated and not as a part of ear-voice span. Here, too, paraphrase was taken into account, and there were specific rules relating to the number of words which could have been meaningfully translated into English. For instance, ne . . . pas was counted as one word from the ne, and articles in the original were not counted when they would not normally have been translated.

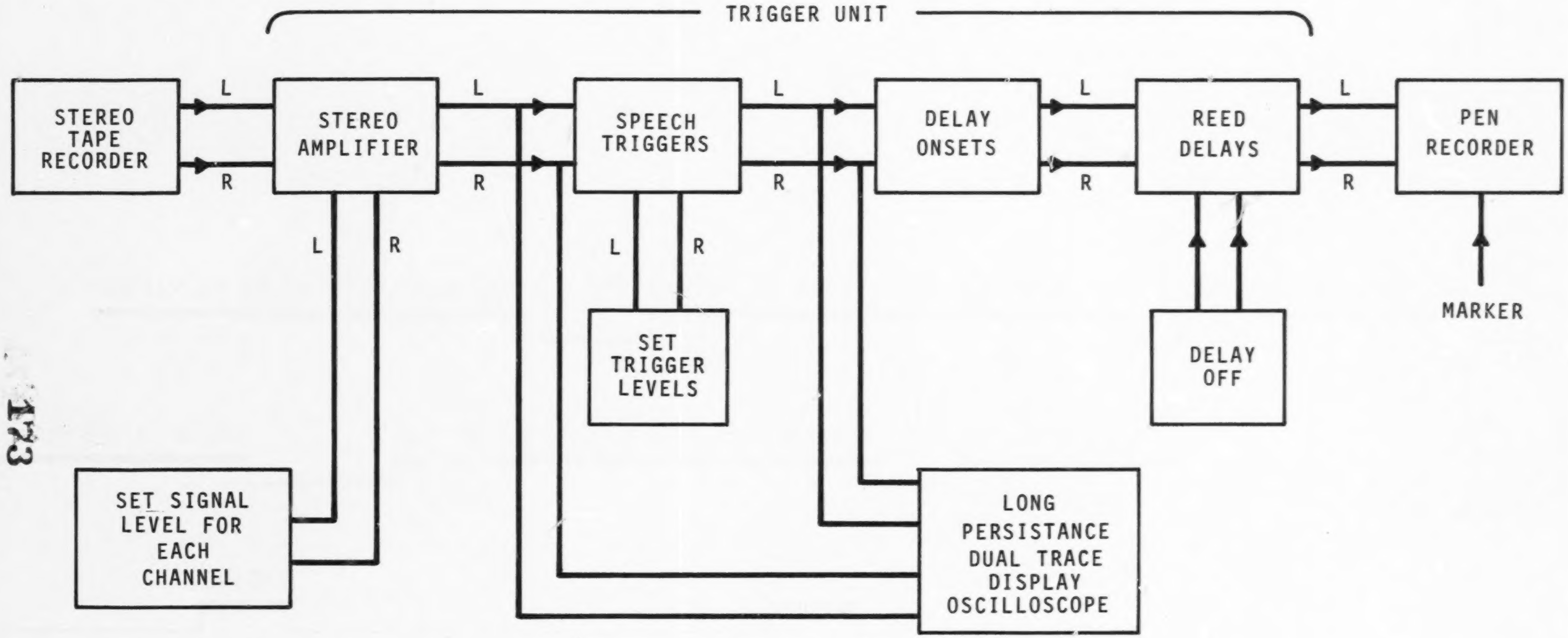

TRIGGER UNIT
STEREO TAPE RECORDER
STEREO AMPLIFIER
SPEECH TRIGGERS
DELAY ONSETS
REED DELAYS
PEN RECORDER
SET SIGNAL LEVEL FOR EACH CHANNEL
SET TRIGGER LEVELS
LONG PERSISTANCE DUAL TRACE DISPLAY OSCILLOSCOPE
DELAY OFF
MARKER
L
R

Figure 15.1. Block diagram of trigger unit input and output.

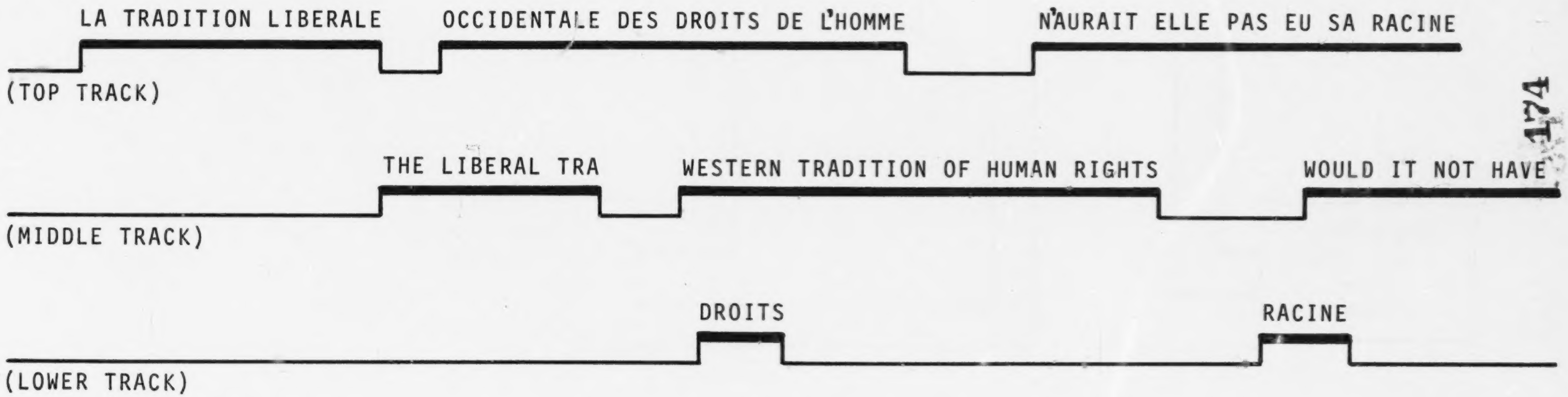

LA TRADITION LIBERALE
OCCIDENTALE DES DROITS DE L'HOMME
N'AURAIT ELLE PAS EU SA RACINE
(TOP TRACK)
THE LIBERAL TRA
WESTERN TRADITION OF HUMAN RIGHTS
WOULD IT NOT HAVE
(MIDDLE TRACK)
DROITS
RACINE
(LOWER TRACK)

Figure 15.2. Pen tracings of source tape (top track), interpreter's version (middle track), and markers (lower track).

Results

Analyses of variance showed the following main and interaction effects:

1. Words correctly shadowed or translated. Significantly more words were correctly shadowed than translated (F = 6.767; df = 1, 8; p < .05), and the effect of presentation rate was also significant (F = 24.752; df = 4, 32; p < .001). As can be seen from Figure 15.3, the significant interaction (F = 4.363; df = 4, 32; p < .05) indicated that presentation rate had a greater effect on the performance of the interpreters than of the shadowers.
2. Ear-voice span. As can be seen from Figure 15.4, interpreters had significantly greater ear-voice spans than shadowers (F = 56.304; df = 1, 8; p < .001), and there was a significant effect of presentation rate (F = 11.408; df = 4, 32; p < .001). The significant interaction also indicated that presentation rate had a greater effect on interpreters (F = 5.029; df = 4, 32; p < .001).
3. Number of utterances. Figure 15.5 shows that interpreters tended to produce more utterances (i.e., pause more often) than shadowers (F = 3.742; df = 1, 8; p < .05), and that both groups produced significantly fewer utterances (paused less often) as presentation rate increased (F = 3.377; df = 4, 32; p < .05). The interaction term (F = 0.064) did not approach significance.
4. Number of words per utterance. Figure 15.6 appears to show that shadowers produced more words per utterance than interpreters, and that there is a significant interaction between presentation rate and words per utterance, but these results do not approach significance on the analysis of variance.
5. Time per utterance. As can be seen from Figure 15.7, input rate had differential effects on the mean utterance times of shadowers and interpreters (F = 3.09; df = 4, 32; p < .05). The effect of presentation rate was significant (F = 8.57; df = 4, 32; p < .001), but there was no significant difference between groups.
6. Pause times. Though Figure 15.8 appears to show definite main and interaction effects, these results are not significant.
7. The ratio of total pause time to total speech time. On the average, interpreters maintained higher pause-speech ratios than shadowers (F = 8.813; df = 1, 8; p < .001). Rate of presentation had a significant effect on pause-speech ratios for both groups (F = 3.303; df = 4, 32; p < .05), but the interaction term was not significant. These results are illustrated in Figure 15.9.
8. Output rate. As can be seen from Figure 15.10, shadowers maintained a consistently higher rate of output than interpreters (F = 9.983; df = 1, 8; p < .05). Presentation rate had a significant effect on the performance of both groups (F = 4.697; df = 4, 32; p < .05), and the interaction term was also significant (F = 4.363; df = 4, 32; p < .01).
9. Output variability. The correlations between words per utterance and time per utterance were calculated for Ss in each group at each presentation rate, and the weighted average correlations were then calculated

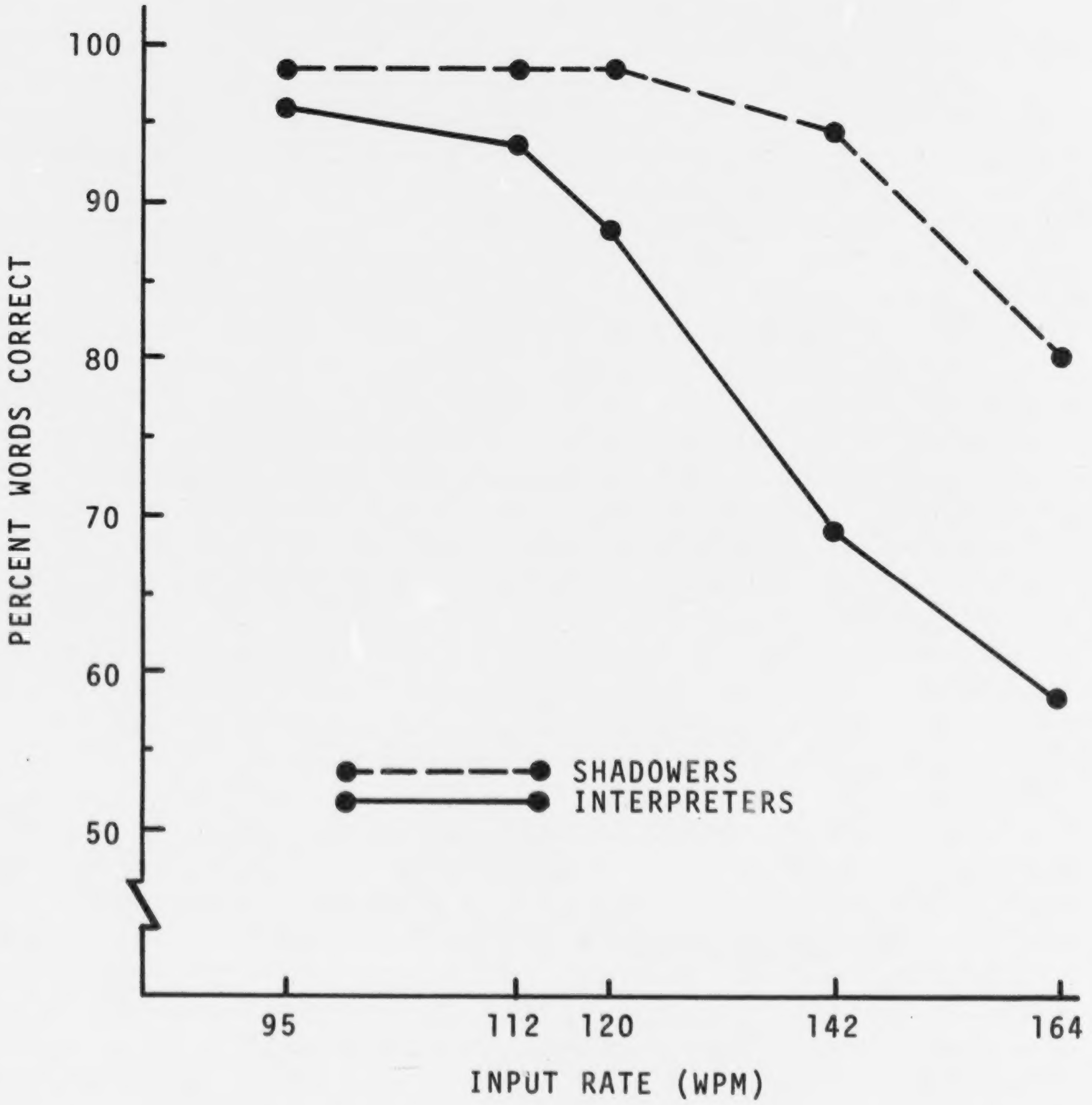


Figure 15.3. Mean % words correct.

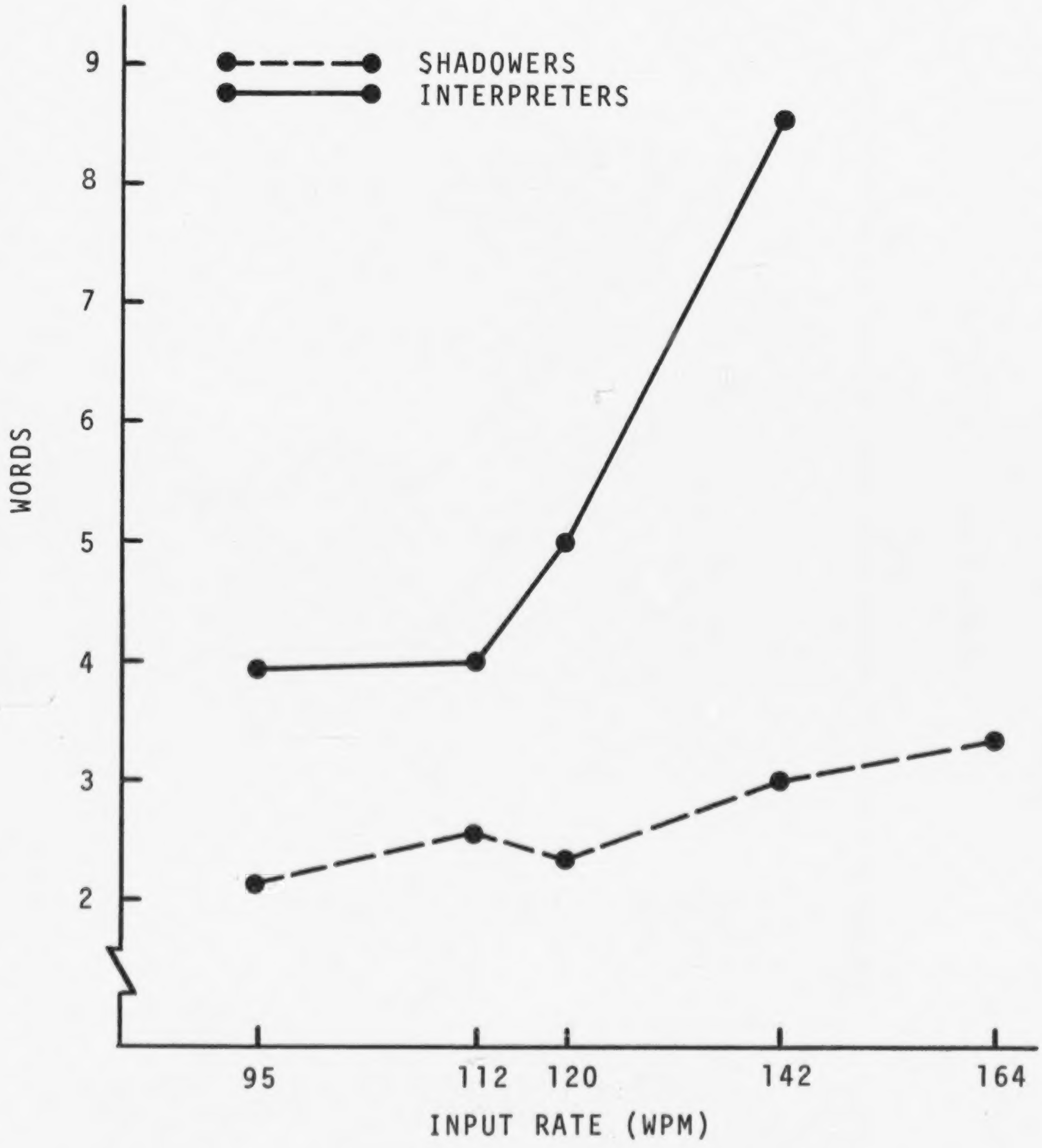


Figure 15.4. Mean ear-voice span.

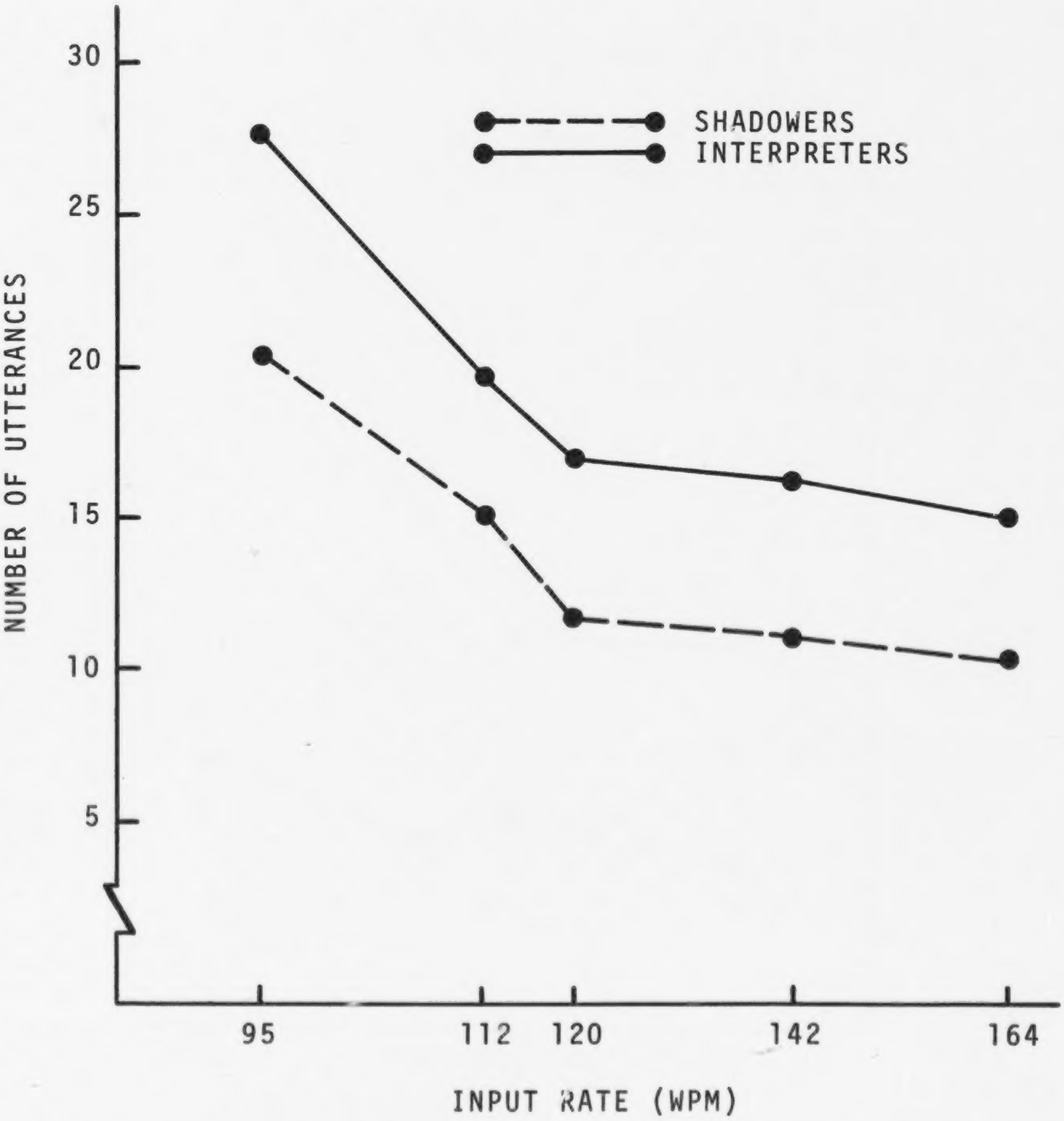


Figure 15.5. Mean number of utterances.

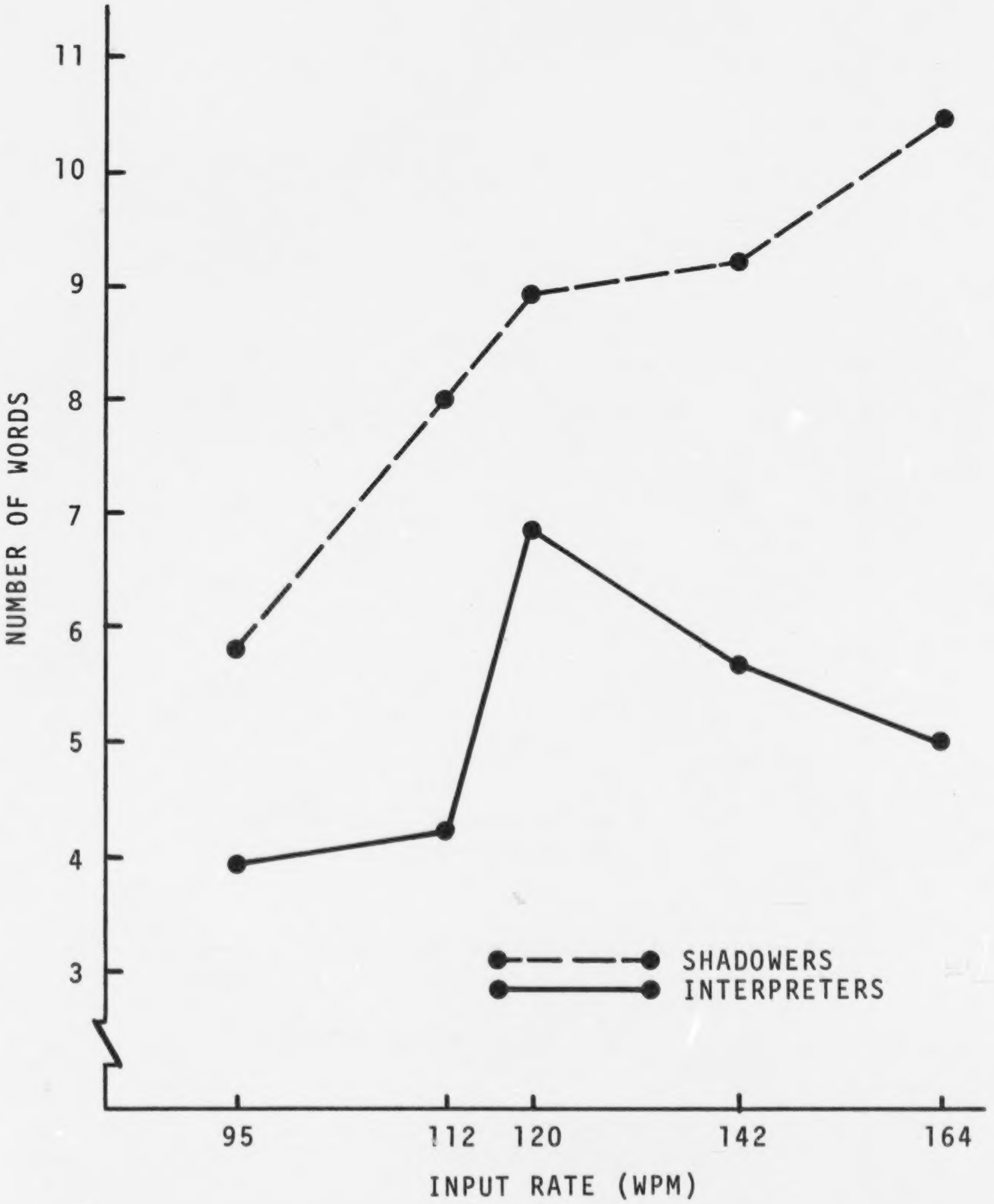


Figure 15.6. Mean number of words per utterance

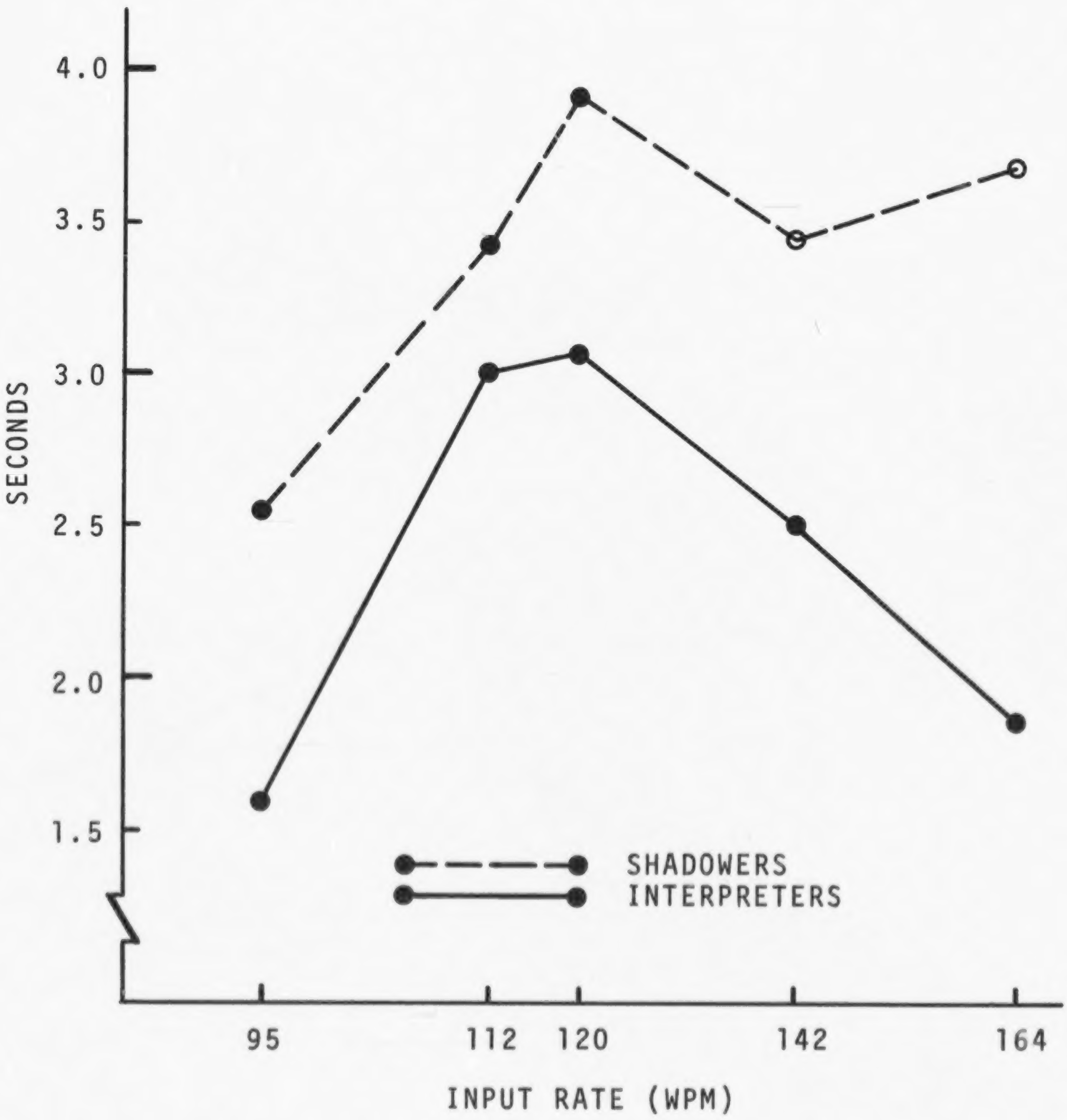


Figure 15. 7. Mean utterance time (secs.).

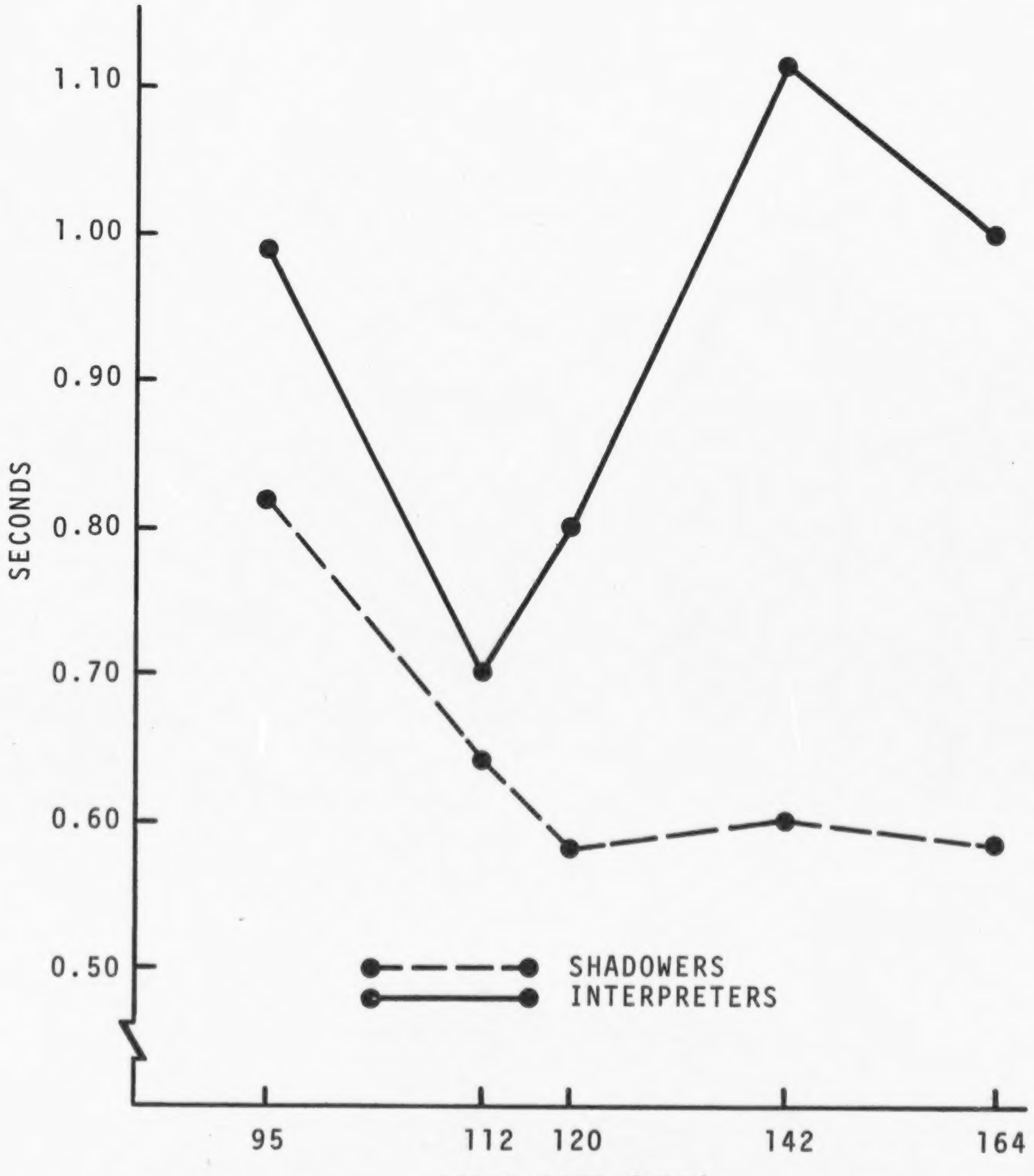


Figure 15.8. Mean pause times (secs.).

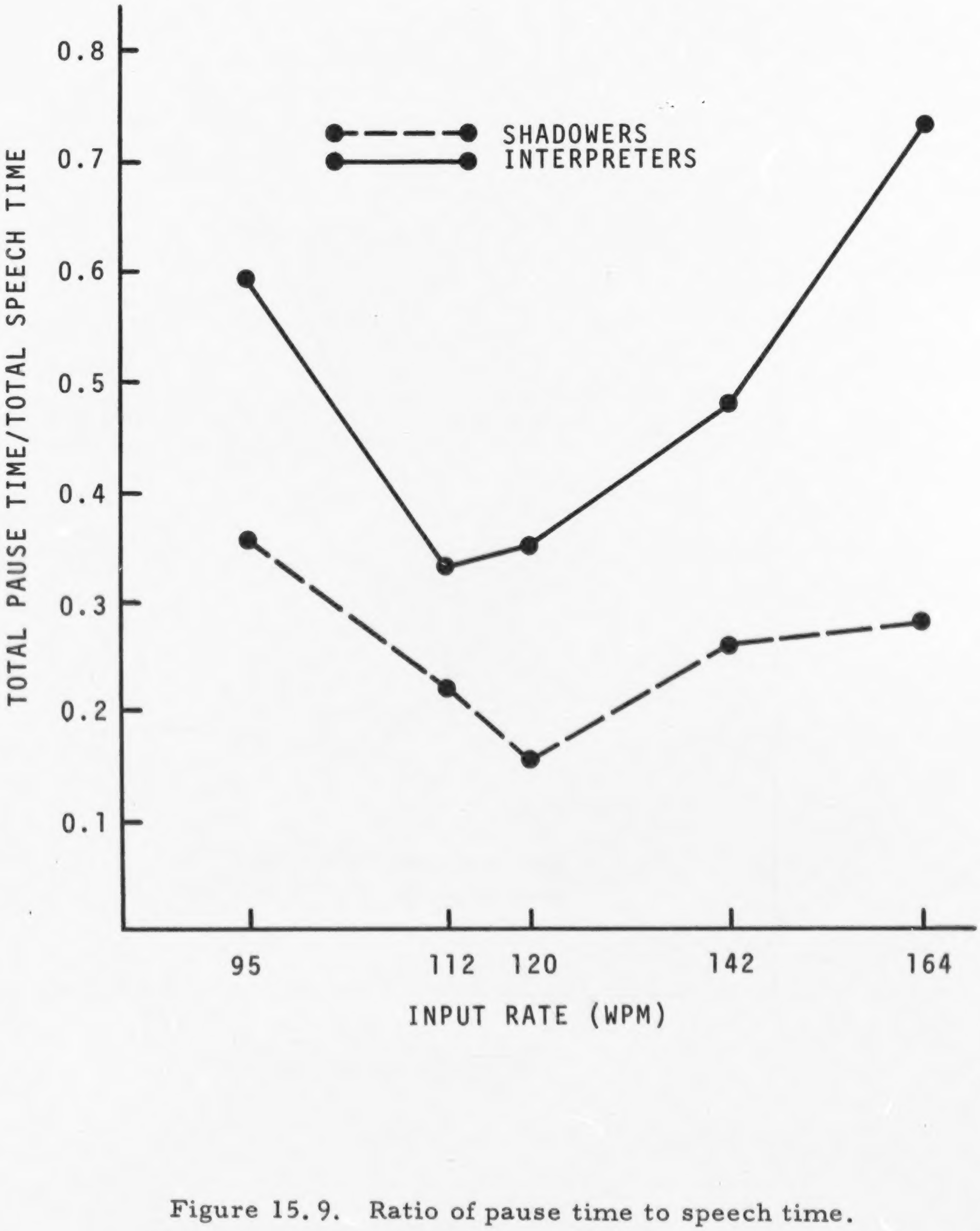


Figure 15. 9. Ratio of pause time to speech time.

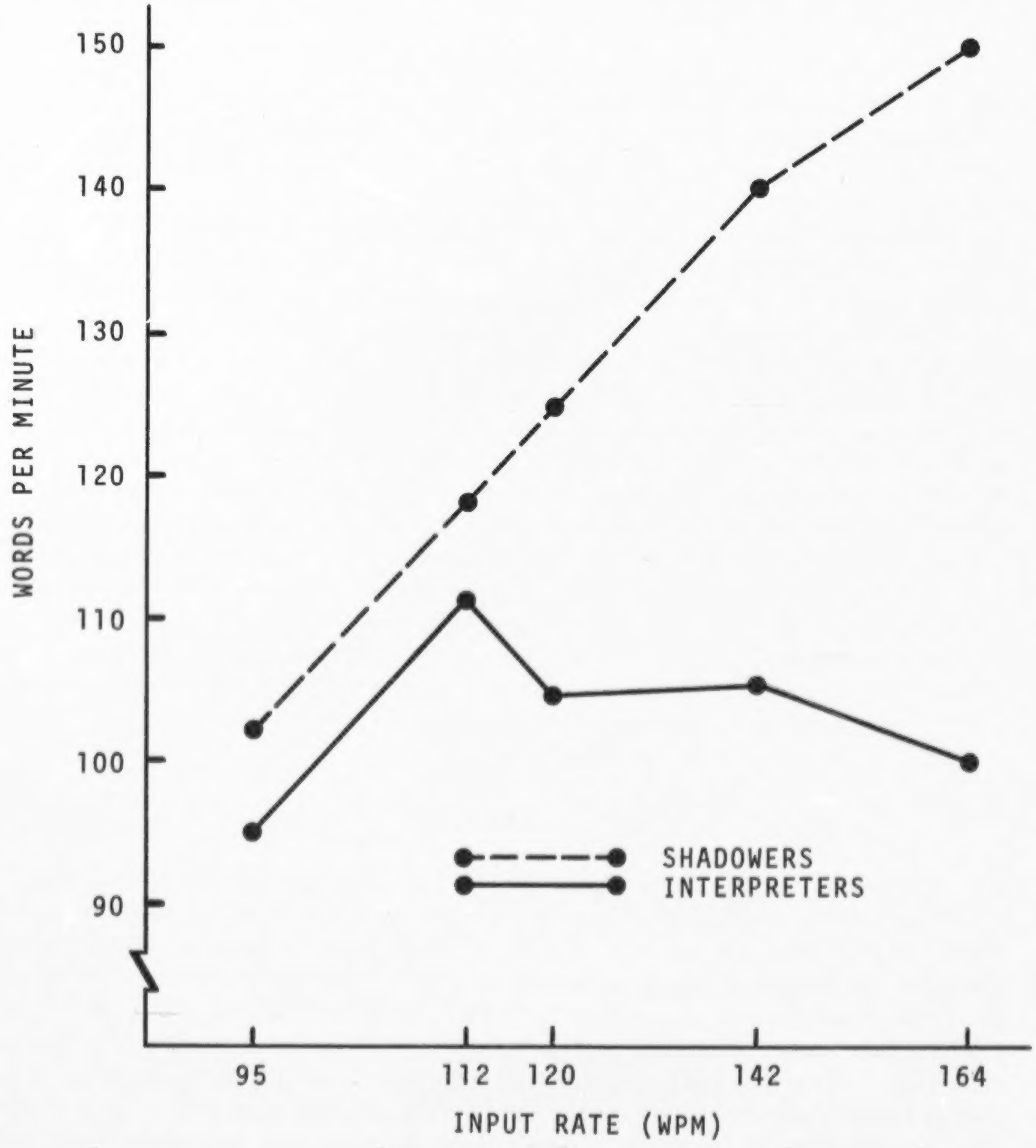


Figure 15.10. Mean output rates (wpm).

(McNemar, 1962). The plot of these relationships in Figure 15.11 shows that both groups became less variable in terms of their output rate per utterance as presentation rate increased.

Discontinuities in Output

The frequencies within each category are shown in Table 15.1. Since these categories are not independant (i.e., the number of omissions affects the number of substitutions and corrections a S can make), the data are not suitable for an overall statistical analysis, but some tentative conclusions may still be drawn from the results.

1. From the totals for groups and rates, it can be seen that interpreters tended to produce more discontinuities than shadowers, and that these increased with increase in presentation rate.
2. Both groups omitted a similar number of single words, but interpreters omitted more phrases and longer passages. Here, too, there was an effect of presentation rate. The average number of words in phrases omitted by interpreters and shadowers was 3.7 and 4.3 respectively. The average length of longer omissions by interpreters was 15.2 words.
3. While shadowers substituted more single words than interpreters, the latter substituted more phrases. The average length of phrase substitutions was 4 and 3.3 words for interpreters and shadowers respectively.
4. Interpreters corrected more single words and markedly more phrases than shadowers. The average length of phrase corrections was 2.8 words for both groups. It was interesting to note that no words were omitted after corrections by shadowers, and that there were only four omissions after corrections (of phrases) by interpreters.

Discussion

The fact that significantly more words were correctly shadowed than were correctly interpreted suggests that any decrement in interpreters' performance was due to the effects of presentation rate on the processes involved in interpretation rather than to an inability to perceive and repeat the input message correctly. The results confirm Carey's (1968) finding that fewer words are correctly shadowed at the faster presentation rate. Carey's fastest rate was 180 wpm, while the fastest rate in the present study was 164 wpm. It is only at this last rate, however, that shadowers' performance deteriorates, whereas interpreters' performance falls off with each increase in rate. Both shadowers and interpreters had less time in which to perceive and speak at faster rates, but interpreters also had less time in which to decode from French and encode into English. Unlike the interpreter, the shadower only has to repeat, not to understand, what he hears. In a sense, shadowers' scores (words correct) are similar to intelligibility test scores, while interpreters' scores demonstrate both intelligibility and

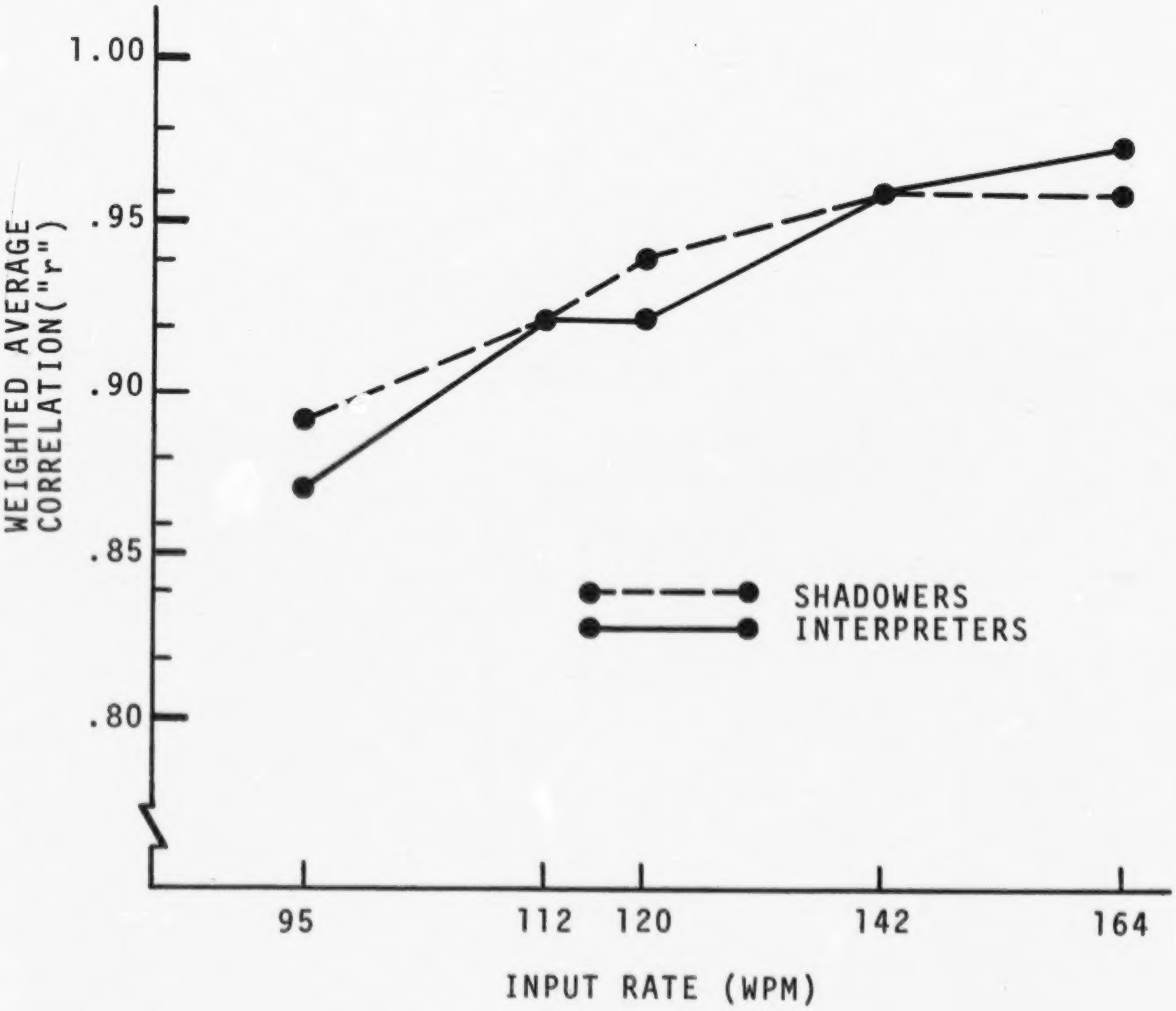


Figure 15.11. Weighted average correlations words per utterance/ time per utterance.

TABLE 15.1

FREQUENCIES OF DISCONTINUITIES

Input rate wpm	Omissions				Long Omissions		Substitutions				Corrections				Total	
	Word		Phrase				Word		Phrase		Word		Phrase			
	Shad.	Int.	Shad.	Int.	Shad.	Int.	Shad.	Int.	Shad.	Int.	Shad.	Int.	Shad.	Int.	Shad.	Int.
95	1	5	1	12	0	0	5	4	3	5	3	2	2	11	15	39
112	7	9	1	6	0	0	8	7	1	5	2	5	0	2	19	34
120	2	9	2	12	0	0	6	6	1	7	1	4	0	3	12	39
142	4	6	6	20	0	4	15	6	5	8	2	4	0	4	32	52
164	20	7	22	22	0	9	26	11	6	15	3	6	1	5	78	76
Total	34	36	32	72	0	13	60	34	16	40	11	21	3	25	156	240

comprehension. As Foulke and Sticht (1967a) have demonstrated, presentation rate has a greater effect on comprehension than on intelligibility, and it is to be expected, therefore, that interpreters should make more errors than shadowers at faster input rates.

As would be expected from Treisman's (1965) experiment, the ear-voice span was greater for interpreters than for shadowers. Though shadowers' ear-voice spans rose only slightly from slowest to fastest presentation rate, the interpreters' ear-voice spans almost doubled over the same range. When these results are considered together with Ss' output rates at each presentation rate, it can be seen that shadowers were able to increase their output rates as input rate increased at the cost of only a slight increase in ear-voice span. Interpreters, however, seemed only able to maintain fairly steady output at the expense of lagging further and further behind as input rate increased. There would seem then to be an optimal output rate for interpreters, and in order to maintain it in the face of faster input rates, they were forced to lag further and further behind.

Ear-voice span, whether for shadowing or interpreting, is attributable to the accumulation of items in some form of short-term buffer store, while previously received information is processed by a central mechanism. When shadowing, this process probably involves analysis of the auditory input at the level of the phoneme, syllable, or word, and direct recoding in terms of the articulatory movements required to produce the sounds just heard. In interpretation, however, more complex analysis of the input message must be carried out, and larger grammatical units must be involved in order to derive the deep structure of the source-language message and translate to the surface structure and phonetic output of the target language. Since translation must involve larger "units" than shadowing, it seems reasonable to suppose that the major constituent or phrase might be the minimal unit of analysis in interpretation. One would, therefore, expect ear-voice span to be greater not only because processing takes longer, but because the constituent may also be the unit of storage.

It was hypothesized that interpreters would either reduce pause length and increase utterance length as input rate increased, or that they would pause more frequently and speak for shorter periods. A further prediction was that their output rate would become less variable as presentation rate increased. While only the last prediction has been supported by the results, it is worth noting that though shadowers were able to redistribute their performance in time in the manner predicted for interpreters, the latter were able to optimize their use of speech and pause times by speaking more and pausing less up to a presentation rate of 120 wpm, but then began to pause more and speak less. At input rates of 120 wpm and over, the interpreters were lagging further and further behind and making more and more errors. They spoke at a steady rate but only after longer pauses.

If, as was suggested above, shadowing involves a comparatively low level of processing, then it is not surprising that processing rate can keep track with input rate, at least within the range of input rates employed in this study. So long as the shadower can keep fairly close to the input he will need neither to utilise input pauses (where possible) nor to make extra pauses himself in order to process a backlog of material. The interpreter, having to cope with larger units before being able to translate, finds that as the intervals between items (words, phrases) become shorter than the time taken to process them, he must effectively slow down the rate at which he works. Finding that he cannot increase his own overall rates of processing and output, he appears to opt for a strategy of working in bursts and must lengthen pause times in order to do so. The extra time thus made available should enable him to cope with the increasing backlog of material in short-term store, but items in store accumulate and deteriorate faster than the interpreter can cope and, in fact, his performance falls off.

The principal effect of increasing presentation rate was to increase the number of discontinuities in all categories. Carey (1968), in order to account for his Ss increases in errors in shadowing at faster presentation rates concluded: ". . . when a speaking error is made, and monitoring indicates that what was spoken does not agree with the input, the mismatch may demand additional time that could have been devoted to perceiving the next section of input. Once speaking errors begin the result is a snowballing effect that results in a long stretch of omitted words." Contrary to this suggestion, however, the shadowers in the present study did not omit long stretches of words at faster input rates. Shadowers omitted mainly smaller units, whereas interpreters omitted more, and longer, passages at faster presentation rates.

At faster presentation rates the responses of both groups became less precise, as indicated by the larger numbers of word and "phrase" substitutions. It is worth noting that shadowers tended to substitute more single word units, whereas interpreters substituted more "phrases." These results, together with the fact that shadowers tended to correct words rather than "phrases," whereas interpreters omitted, substituted, and corrected "phrases" rather than words, suggests that interpreters do indeed work with larger units than shadowers. Though no attempt has been made here to analyze the structure of the "phrases" omitted, substituted, or corrected, that is to say whether they involve major or minor constituents, or whether between constituent discontinuities also occur, further research on these lines might help to answer the question as to the "unit" of storage, analysis, and monitoring in this complex information processing task.

The very fact that interpreters did correct their own output demonstrates that they do monitor what they say. These corrections are usually corrections of previous substitutions but may also be improvements or changes of already acceptable translations. In spite of the complexity of the operations involved, it appears to be possible for Ss to store input whilst both translating and monitoring their own output, without losing input either through interference or trace decay. Out of a total of 25 phrase corrections made by interpreters, only four were followed by omissions. It seems extremely unlikely that active rehearsal of Message 2 can take place while Message 1 is being translated, monitored, and corrected; and unless one postulated that attention can be switched rapidly between these operations, one is led to conclude that attention can be shared between input, translation, and monitoring. The difficulty with an attention switching model lies in the specification of the rate at which switching can occur and also of the duration of each switch. As Moray (1969) points out, the rate at which switching can occur depends upon the unit of analysis, and the larger the unit, the longer the duration. If, as seems likely, the unit of analysis in interpretation is the constituent, then the duration of the switch will be comparatively long. Even if this is between one-half and 1 second, one must still ask what will happen to input arriving while S is either interpreting or monitoring. The present evidence suggests that such information need not be lost, and an attention sharing model seems most likely. At any rate, simultaneous interpretation appears to be a practical situation in which the processes associated with short-term storage need not involve covert repetition.

Conclusions

The aim of this study was not only to examine the effects of message presentation rate upon the simultaneous interpreter's performance over time, but also to study his output for cues as to the processes involved in so complex a skill.

If Ss have only to shadow a continuous message they are able to keep up with faster presentation rates by speaking more quickly, lengthening their utterances, and shortening pauses between utterances. When required to simultaneously interpret a message into a target language, faster input rates cause Ss to lag further and further behind and to make more errors than shadowers. In order to maintain a steady output rate, these Ss pause more and speak in shorter bursts. Though both shadowers and interpreters correct their errors, interpreters tend to work in units of 2-3 words or more. This, together with evidence from omissions and substitutions, suggests that the unit of analysis is the "phrase" for interpreters (where understanding is required) and the word for shadowers (where S is required to demonstrate perception rather than comprehension).

The picture emerges of an information handling system which is subject to overload if required to carry out more complex processes at too fast a rate and copes with overload by reaching a steady state of throughput at the expense of an increase in errors and omissions. There is evidence that attention is shared within this system between the input message, processes involved in translating a previous message, and the monitoring of feedback from current output. Under normal conditions, attention can be shared between these processes, but when the total capacity of the system is exceeded, less attention can be paid to either input or output if interpretation is to proceed at all. Hence, less material is available for recall for translation, and more omissions and uncorrected errors in output will occur.

CHAPTER XVI

THE A. F. B. HARMONIC COMPRESSOR

John W. Breuel and Leo M. Levens*

Background Information

We have seen an ever increasing volume of publication of information of all sorts. Much of this material is being converted to various recorded forms. There is a great need for compacting the bulk of this material and reducing the time required for reading it.

A previous paper presented at the October 1968 Audio Engineering Society Convention outlined methods of bulk reduction. This paper presents a method of time compression called Harmonic Speech Compression. The general definition of speech compression refers to any method of reducing the time required to transmit a spoken message.

Many of the readers of the Talking Books for the Blind found the normal pace of reading too slow. Some people were not accomplishing all the reading they required. Others lost interest due to the slow speed of many of the recordings. The 45 rpm turntables, when they became available in the late 1940's, provided a partial answer. The 33 1/3 rpm disc could be played at 45 rpm. More reading was accomplished, interest was heightened, and comprehension did not suffer.

The Books for the Blind program has grown to the point where it is impossible for any person to read more than a small fraction of the daily output. Many blind people derive their greatest satisfaction from aural reading. Listening to talking books, magazines, newspapers, radio, and television provides their chief source of information. Blind students use tape and disc recordings instead of or in addition to braille reading. This pointed up the need for accelerating aural reading.

The Foundation undertook the development of variable frequency power supplies to increase the speed of the induction or synchronous motors in tape and disc reproducers. This was a partial solution. Comprehension did not suffer in most cases until 65% to 75% overspeed. At that point the extreme distortion made the recording unintelligible to most people.

---

*John W. Breuel and Leo M. Levens are with the American Foundation for the Blind, 15 West 16th Street, New York, New York 10011. The paper was presented by Mr. Levens.

Dr. Grant Fairbanks of the University of Illinois developed another method of speech compression in the early 1950's. This method did not result in pitch distortion. Speech on tape was sampled in small segments, some segments were discarded, and the consequent gaps were closed. This method provided some additional speech compression capability but presented some limitations. Among these were: an annoying tone due to the discard rate, a limited input and output format, and a constant maintenance problem in the rotating magnetic pickup head and brushes.

Many of the people working in the field thought that an electronic means of accomplishing speech compression would present significant technological advantages.

These advantages would be: the ability to process at a high speed (such as in multiple speed duplication), multichannel processing (for multitrack tapes), greatly improved speech quality by eliminating the discard rate noise, and simpler and more flexible operation and maintenance.

Initially the Foundation's Engineering Division considered dividing the accelerated voice spectrum into a number of frequency bands and converting these frequencies to lower frequencies by modulation. It was found that this method destroyed the harmonic relationship of the voice frequency components and produced severe distortion. We became aware of Bell Telephone Laboratories' work on bandwidth reduction by means of harmonic compression. This method reduced frequencies by division in contrast to the subtraction involved in modulation.

No harmonic speech compressor had ever been built. The system had been simulated on a digital computer. Short samples of speech were of excellent quality. Drs. M. R. Schroeder and R. M. Golden of Bell Laboratories provided the Foundation with a block diagram of a system suitable for our use. Bell Laboratories provided specifications for the filters.

The Foundation undertook development and construction of the circuits and system in 1965. We provided the production engineering for the filters and the development and construction of the electronic circuitry. The experimental prototype system was completed in 1968. The results of the experiment bore out the promises of the computer simulation.

Harmonic Compressor System and Circuits

Figure 16.1 is a block diagram of the complete harmonic compressor system.

The 36 channel filter band separates the voiced speech into its individual harmonically related frequency components. The individual output frequencies may be considered carriers with narrow band amplitude and phase modulations. Their bandwidths are proportional to the syllabic rate rather than to the fundamental pitch frequency. The filters are Bessel bandpass filters with bandwidths of less than 100 Hz (at normal speed). The filter skirts have steep slopes. The frequency response of the filter bank is flat within 1 db and has linear phase response.

Each filter output feeds its corresponding frequency divider. The divider preserves the original amplitude and phase relationships over a range of 70 db. A block diagram of one of the frequency divider circuits is shown in Figure 16.2.

Each frequency divider circuit receives an input consisting of one of the harmonically related frequencies from one of the bandpass filters. This signal is applied to an amplifier and simultaneously to a zero crossing detector. Waveforms of the input and outputs of the audio amplifier and the zero crossing detector are shown in Figure 16.3. Zero crossings are detected by "infinite clipping." The clipped wave is shaped into a trigger pulse. This pulse controls a bistable multivibrator.

The outputs from both multivibrator transistor collectors are used to drive the gates of two Metal Oxide Semiconductor Field Effect Transistor (MOSFET) chopper modulators in opposite phases. The two opposing phase outputs of the audio amplifier phase inverter are fed to the drains of the MOSFET chopper modulators. The two modulator outputs are combined by a summing amplifier and passed through a low pass filter to remove any distortion.

Modulator and output waveforms are shown in Figure 16.4.

The output from the summing amplifier is a sine wave of the original frequency with every other period inverted. This is a waveform of half the original frequency with an amplitude and phase proportional to the original frequency. The original frequency appears as a second harmonic of the new divided frequency and is removed by subsequent output filtering.

The outputs from all the channels now appear as undistorted sine waves of one-half their respective input frequencies.

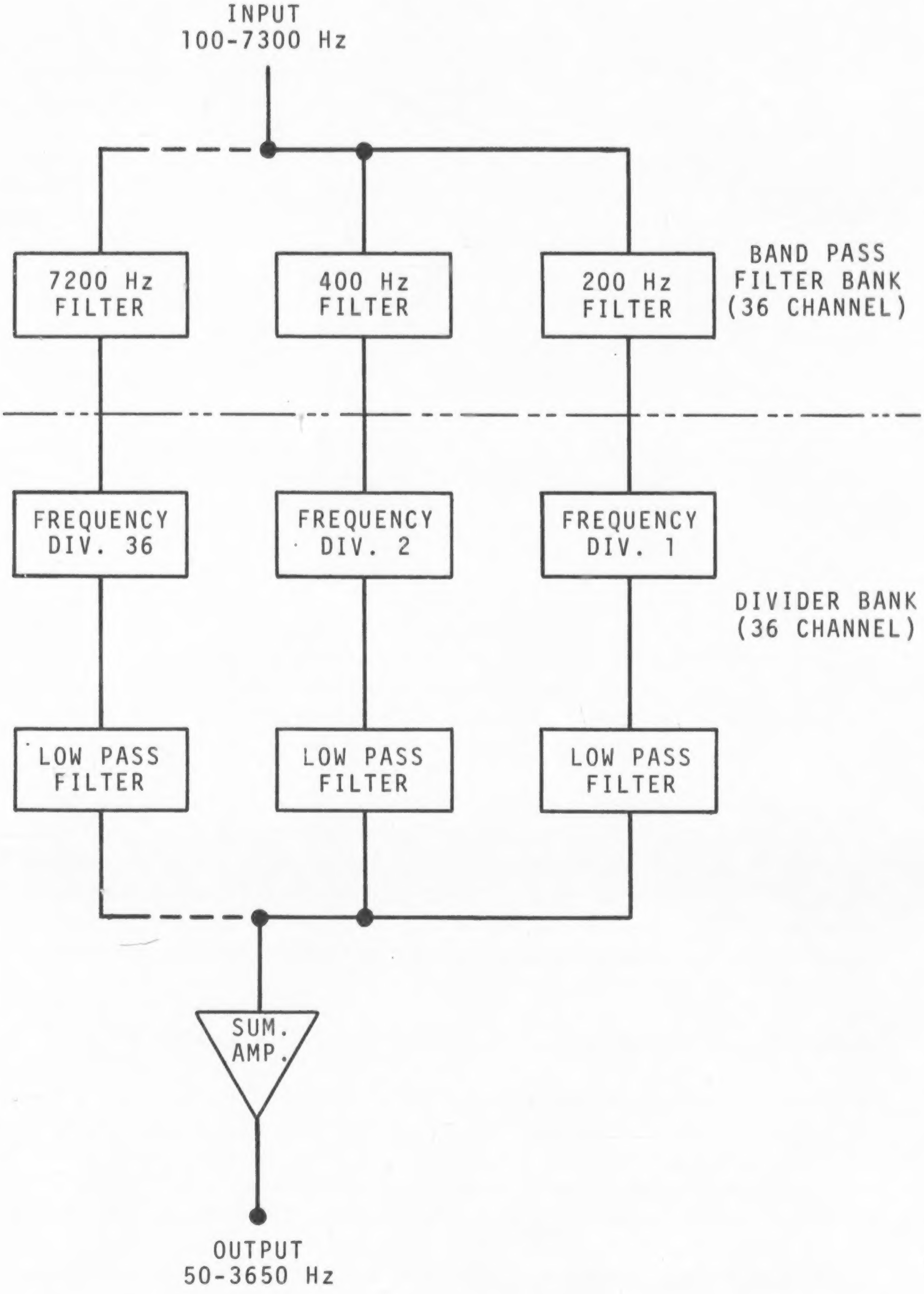

INPUT
100-7300 Hz
7200 Hz FILTER
400 Hz FILTER
200 Hz FILTER
BAND PASS FILTER BANK (36 CHANNEL)
FREQUENCY DIV. 36
FREQUENCY DIV. 2
FREQUENCY DIV. 1
DIVIDER BANK (36 CHANNEL)
LOW PASS FILTER
LOW PASS FILTER
LOW PASS FILTER
SUM. AMP.
OUTPUT
50-3650 Hz

Figure 16.1. Harmonic compressor system block diagram.

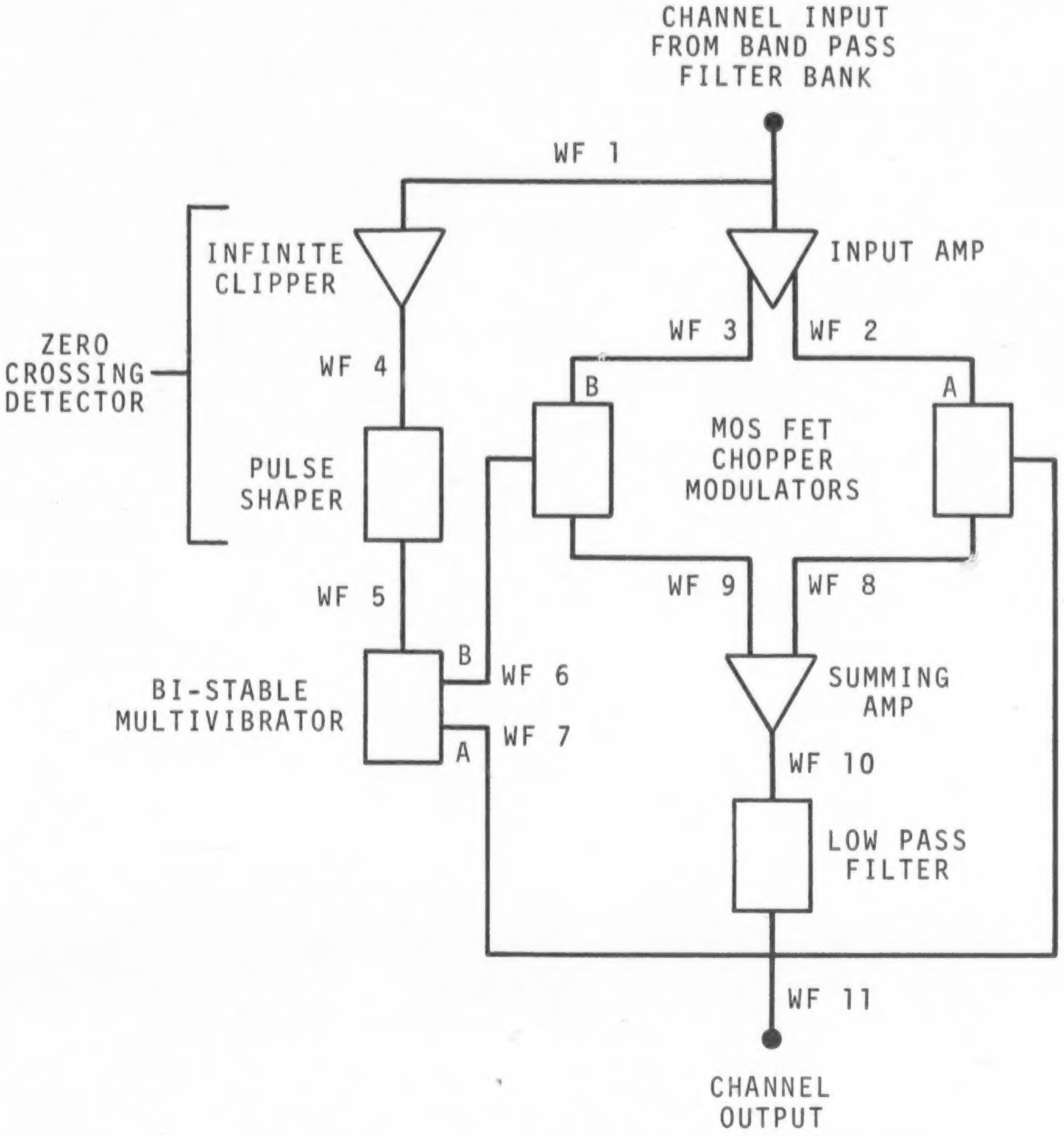

CHANNEL INPUT FROM BAND PASS FILTER BANK
WF 1
INFINITE CLIPPER
INPUT AMP
ZERO CROSSING DETECTOR
WF 4
WF 3
WF 2
B
A
MOS FET CHOPPER MODULATORS
PULSE SHAPER
WF 5
WF 9
WF 8
B
WF 6
BI-STABLE MULTIVIBRATOR
WF 7
A
SUMMING AMP
WF 10
LOW PASS FILTER
WF 11
CHANNEL OUTPUT

Figure 16.2. Harmonic compressor frequency divider channel diagram. (Waveform numbers refer to Figures 16.3 and 16.4.)

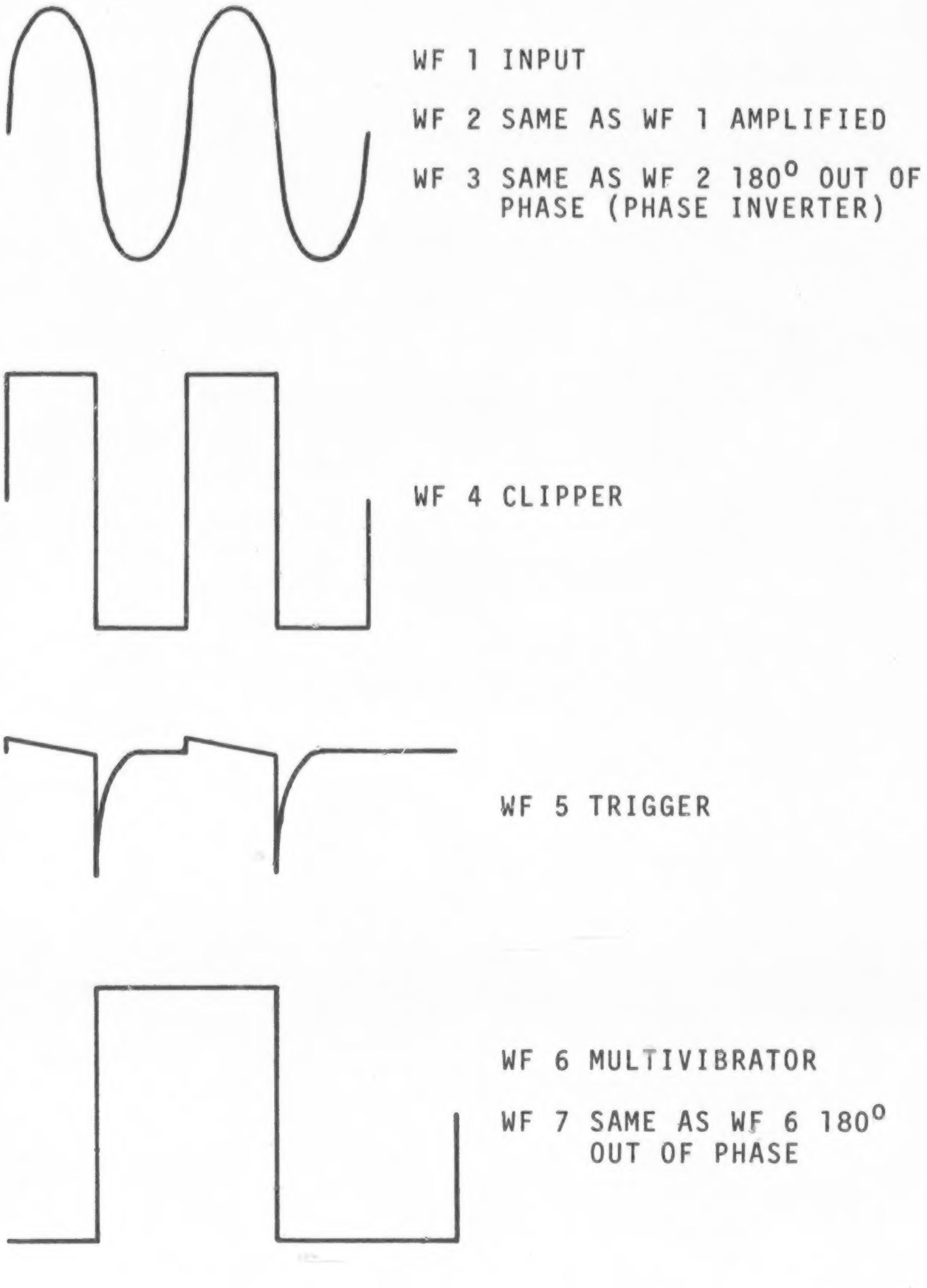

WF 1 INPUT
WF 2 SAME AS WF 1 AMPLIFIED
WF 3 SAME AS WF 2 180° OUT OF PHASE (PHASE INVERTER)
WF 4 CLIPPER
WF 5 TRIGGER
WF 6 MULTIVIBRATOR
WF 7 SAME AS WF 6 180° OUT OF PHASE
WF (WAVE FORM)

Figure 16.3. Harmonic compressor waveform diagrams. (See Figure 16.2 for waveform locations.)

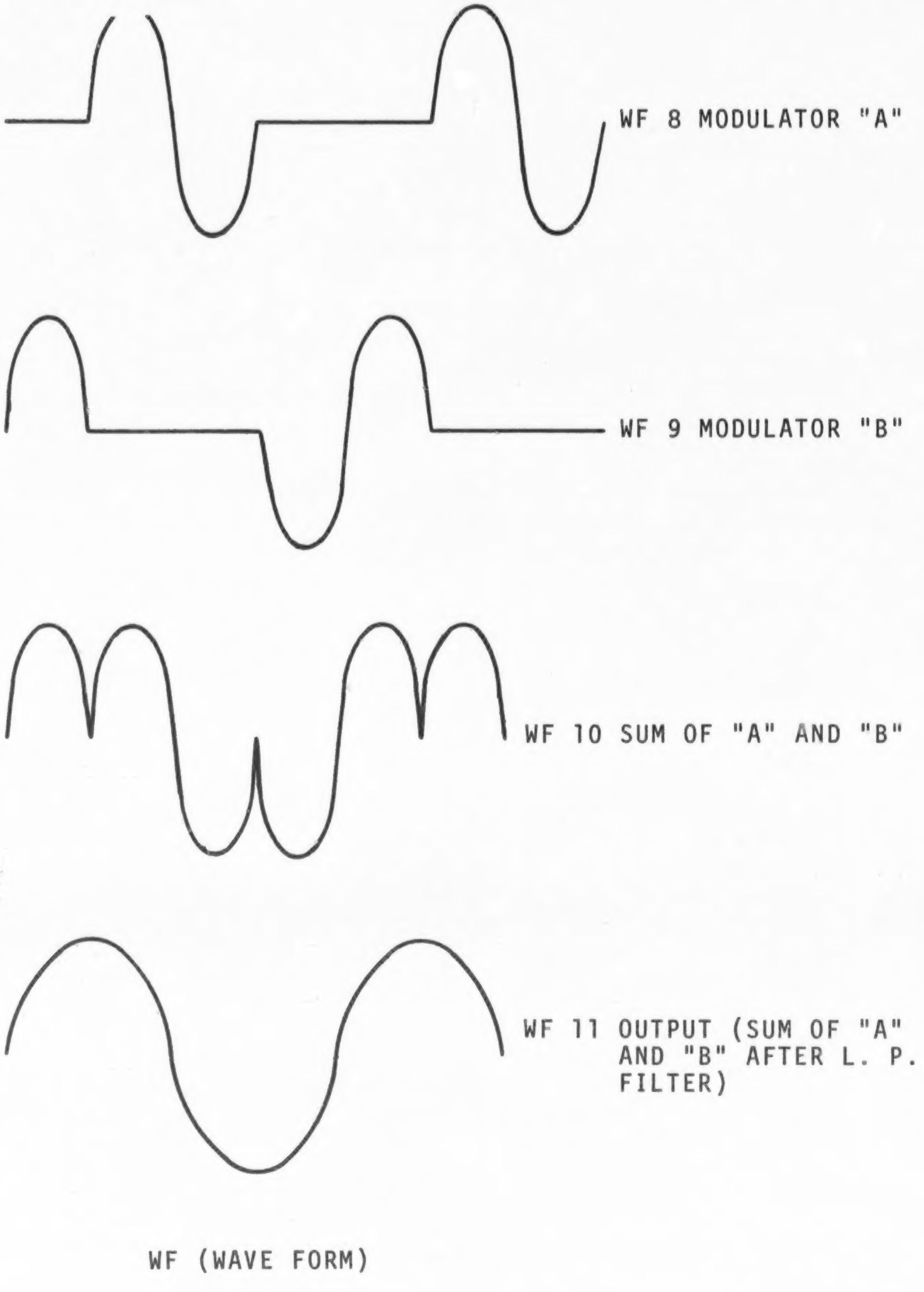

WF 8 MODULATOR "A"
WF 9 MODULATOR "B"
WF 10 SUM OF "A" AND "B"
WF 11 OUTPUT (SUM OF "A" AND "B" AFTER L. P. FILTER)
WF (WAVE FORM)

Figure 16.4. Harmonic compressor MOSFET chopper modulator. (See Figure for waveform locations.)

These outputs are combined in a summing amplifier and fed to a tape or disc recorder.

When an input voice signal is applied at double its normal speed, the output signal is reproduced at double its normal rate with its normal frequency or pitch.

Conclusions

As a result of initial system tests we have reached these conclusions:
1. The system, initially only simulated on a digital computer, operates satisfactorily as an electronic system.
2. There is little significant loss of comprehension at speech rates up to 350 words per minute. Trained individuals may be able to comprehend higher word rates.
3. The fixed output ratio of two times input rate may be varied up to plus or minus 20% without annoying pitch change. This may be accomplished by varying the speed of the tape or disc reproducer.

CHAPTER XVII

THE GRAHAM COMPRESSOR, A TECHNICAL DEVELOPMENT OF THE FAIRBANKS METHOD*

Wayne W. Graham**

As an equipment manufacturer, we have assumed our audience for this paper, with some exceptions, will not have an engineering interest in machine design but rather will expect their equipment to deliver the desired product when required, without their being intimately acquainted with all of the mechanical and electronic details.

However, just as our automobile, commonplace as it has become, still requires that we have some knowledge of what makes it responsive to our demands, with guide lines to be followed if we intend to keep it that way, so must magnetic tape recordings applied to compressed speech be given a certain amount of "tender loving care" through the planning and operational stages.

While Grant Fairbanks et al. (1954) did not hit upon the only method for producing compressed speech, the basic start they gave us has proved out well over the testing period. It remains a flexible, dependable tool whether viewed as a device to be used in the clinic or the speech department of a school system.

The basic design has always made possible individual control over the discard sample size. By taking advantage of miniaturization of magnetic heads and improved magnetic tape, Discerned Sound has extended the range of this control so that discard samples as small as 10 milliseconds (msec.) can be established. It should not be inferred that 10 msec. will always be the desired interval, but the range of the machine probably includes the precise interval needed for the specific job, based on the character of the voice used for the master recording and the percentage of change to be introduced in the altered copy.

---

*Disclosure of technical changes and additions made in the Fairbanks, G., Everitt, W. L., and Jaeger, R. P. (1954) device for time or frequency compression-expansion of speech, made by and incorporated in the machine currently manufactured by Discerned Sound.

**Wayne W. Graham is with Discerned Sound, 4459 Kraft Avenue, North Hollywood, California 91602.

A stock reproduce head in a miniaturized size is available at low cost, fortunately. The machine design has made it routine to use these heads without preselection. Special reproduce heads are not important to our device.

Four reproduce heads are used as a set, and it is important that their individual output or volume be matched to the associated heads. Controls are provided which make it possible to match these heads at any specific frequency.

When they function, these multiple reproduce heads are traveling through a 90 degree arc; and, as each head finishes its pass, it depends on the upcoming head to continue the process. Thus, a transfer from head-to-head is set up under conditions which, when viewed as an audio function, is occurring at the most critical point in the operational chain.

First, we desire to maintain a precise 90 degree arc. More than this will give us a corresponding audio overlap, less will result in a corresponding hole in continuity. It would appear that we need only to plan for the magnetic tape to contact these heads for 90 degrees and secure a mechanical control over the transfer operation. Unfortunately, this is not true. Certain portions of the audio spectrum contained in the recordings (the low frequencies of approximately 200 cycles and below) are much stronger on the tape energy-wise than are higher frequencies, and they will excite the head preparing for its 90 degree pass prior to its establishing intimate contact with the tape. Also, when swinging out of its operating area, it will continue to be responsive to these frequencies even though the head has left the tape. This means, then, that each transfer point presents a moment of time during which two heads are feeding certain frequencies and, since these become combined in the audio chain, they represent a short interval of undesired increased volume.

Discerned Sound design does not accept the limitations of mechanical switching. However, in spite of our efforts, we have not eliminated all of the head transfer disturbance. We have done the following and are continuing to research the problem. We even hope to someday find the magic answer which presently appears unknown to the art. In the meantime, electronics are used to mute each of the multiple reproduce heads except during the time they are in the 90 degree arc. By using what is known as a flip-flop circuit, we have one head or the other, but never two adjacent heads, operating at full level at the same time.

Finally, the reproduce heads are the source of our audio signal, also the source of most of the undesirable noise. The amount of energy picked up from the tape by the head can be read on instruments but is certainly too weak to be usable until amplified many times. Also, since these are rotating heads, a set of slip rings are used to provide a path for the audio signal to leave the moving part and ultimately reach its destination. Without bolstering the strength of the so-called head level signal, any source of noise, such as slip rings, nearby fields of radiation, etc., becomes very difficult to control and can be a constant maintenance problem. Discerned Sound has provided four preamplifiers utilizing hybrid integrated circuits mounted on the rotating fixture which amplifies the signal well above the problem area before it reaches the slip rings.

Another feature of the Fairbanks design, which certainly has been retained intact by Discerned Sound, is the machine's facility for accepting a signal from any usual audio source. This means, for example, that no prerequisite applies to the tape speed which was used at the time of recording the original library. If you have 500 master tapes recorded at 3 3/4 i.p.s. and they are what you wish to use, all you have to do is place them on a tape playback operating at the proper speed and make the altered copy.

A storage bin for the tape, controlled by a pressurized air supply, has made possible an increase in the quantity of tape operating within the machine. This, of course, adds to the convenience and quality of machine performance.

As an equipment manufacturer, we hope we are not visionary in viewing ourselves teamed with you in your efforts to advance compressed speech and its advantages. Our machine represents our entire business interest. We are available, whenever you need us, to discuss what can be done with our unit or consider changes you wish to propose.

CHAPTER XVIII

TIME COMPRESSION OF SPEECH ON A SMALL COMPUTER

S. U. H. Qureshi and Y. J. Kingma*

Introduction

An attempt has been made to find an economical way of compressing speech on a computer and to develop a selective method of compression utilizing the great flexibility offered by a programmed data processor.

The high cost of computer time on large computers is the prohibitive factor in the use of computers for speech compression. The possibility of compressing speech on a small computer, the PDP-8, was therefore investigated, the object being to develop a satisfactory method with a minimum of processing time on the computer.

Early speech research indicated that much of the natural speech signal is redundant (Schroeder, 1966). This redundancy was later demonstrated by Garvey (1953b) who showed that the duration of certain phonemes in normal speech is longer than required for reliable recognition. The object of speech compression is to remove this temporal redundancy and thus convey more information in less time. According to Allen (1967), the ear cannot perceive sounds shorter than 35 milliseconds (msec.) as distinctly separate sounds. Each sound in compressed speech should therefore be at least 35 msec. in duration to be recognizable. We can thus arbitrarily define redundant segments as those parts of speech sounds which exceed 40 msec. in duration and which possess certain features of the speech waveform which do not change appreciably over the whole sound (Endres, 1968). The first step in selective speech compression is to distinguish between various speech sounds. A number of methods of speech segmentation have been developed for the automatic recognition of speech (Gilmour, 1968; Reddy, 1966; Scholtz & Bakis, 1962). The purpose of the segmentation process in this case, however, is to locate redundant parts rather than find sharp boundaries between phonemes. A time-domain method similar to Reddy (1966) is used in preference to others (involving comparison of spectral properties) to avoid costly hardware or excessive computer time in finding the spectrum.

---

*At the time this paper was presented, S. U. H. Qureshi was a graduate student in the Department of Electrical Engineering, University of Alberta, Edmonton, Alberta, Canada, where Dr. Y. J. Kingma is a professor. Mr. Qureshi is now with the Department of Electrical Engineering, University of Toronto, Toronto, Ontario, Canada.

The whole process of compression can be divided into three steps: (1) Input and Feature Extraction, (2) Decision Making, and (3) Removal of Transients and Output.

Input and Feature Extraction

The setup is shown in Figure 18.1. Recorded speech, band limited to 4 kHz, is sampled at 8 kHz, digitized to 12-bit (one sign bit) samples by an A/D converter, and fed into the computer under program control. While the speech samples are temporarily being stored in core and then swapped on to the disc, the program extracts the following three features of the speech waveform for every 12.5 msec. segment:

1. The sound intensity 'I' defined as the absolute maximum of 100 samples where the samples are the ordinates of the speech wave at constant intervals of 125 msec. If the 100 samples are represented by a vector Y then sound intensity

$$I = \max |Y_i| , \quad i = 1, 2, \ldots, 100$$

Y_i being the elements of Y.

2. The waveform asymmetry (Comer, 1966) 'A' defined as the difference between the positive maximum and the negative maximum of 100 samples.

$$A = \max Y_p - \max Y_n$$

Where Y_p includes all the positive elements of Y

Y_n includes all the negative elements of Y

3. The number of zero crossings 'Z' in 100 samples. A zero crossing is said to occur whenever the sign of the i^{th} element of Y is different from the sign of $(i + 1)^{th}$ element i.e., if

$$Y_i \cdot Y_{i+1} = - |Y_i| \; |Y_{i+1}|$$

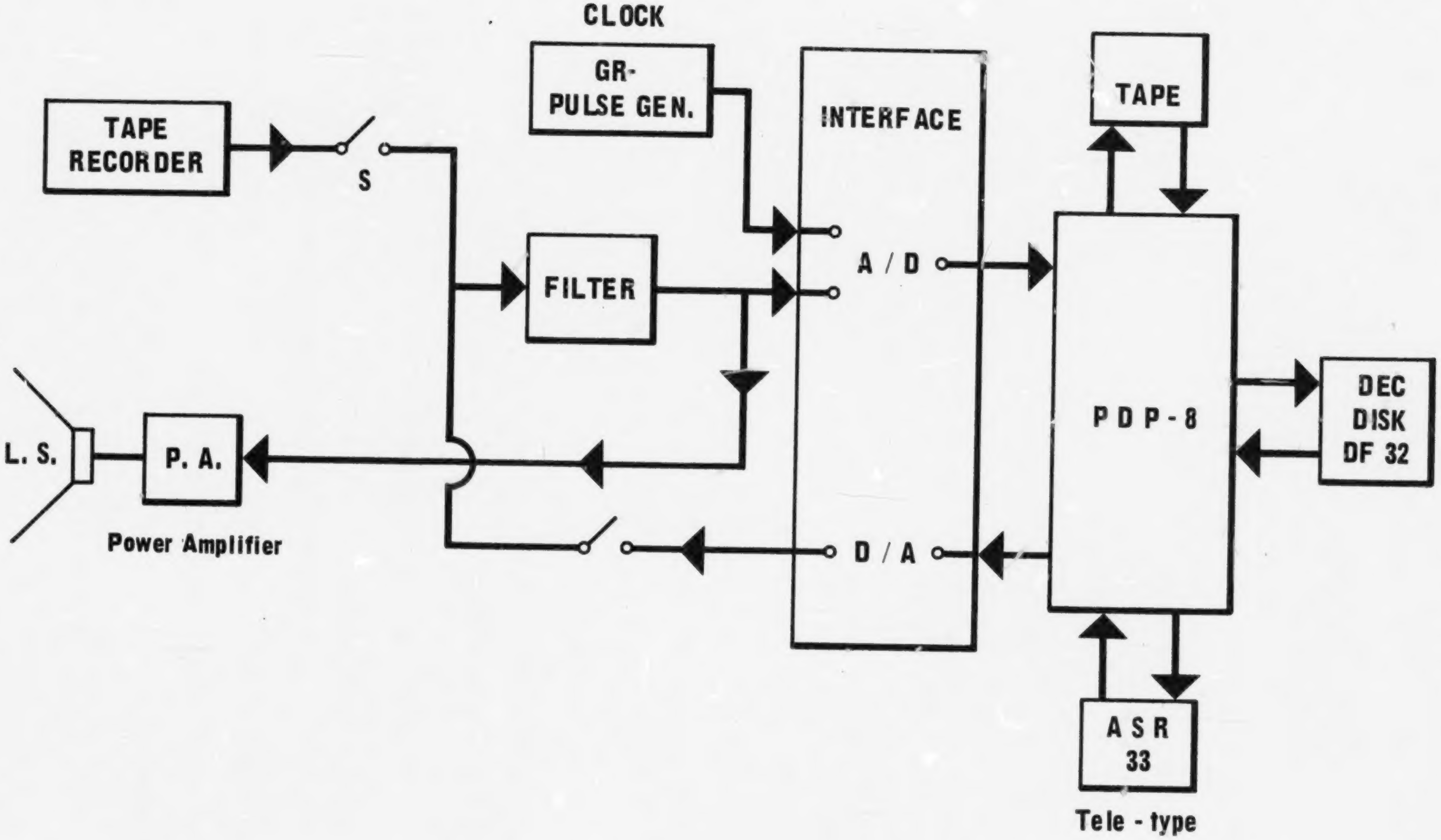

CLOCK
GR-
PULSE GEN.
TAPE
RECORDER
S
FILTER
INTERFACE
A / D
D / A
TAPE
P D P - 8
DEC
DISK
DF 32
A S R
33
Tele - type
L. S.
P. A.
Power Amplifier

Figure 18. 1. System block diagram.

In addition to these three features, the location of the positive peak in each 12.5 msec. segment is stored in the computer core. The locations of the positive peaks are later used to help remove transients that occur due to the deletion of redundant segments.

The speech input continues until the disc memory is full (approximately 3.7 seconds of speech). Control is then transferred to a subroutine which reads the decision making and output programs from the disc into core.

Decision Making

In this program, each of the three features for one segment is compared with those of the contiguous segments in an attempt to find similar segments which can be grouped together to represent a sustained speech sound. The thresholds for decision making (Table 18.1) are set every 125 msec. and depend upon the intensity level of the utterance in the next 500 msec.

TABLE 18.1

THRESHOLDS FOR DECISION MAKING PROGRAM

I_{max} = max intensity in 500 msec. of speech

1.	V/UV threshold (voiced/unvoiced)	6 db.	below	I_{max}
2.	W (threshold for waveform Asymmetry)	18 db.	below	I_{max}
3.	SILENCE threshold	29 db.	below	I_{max}

A flow-chart for the program is shown in Figure 18.2. The similarity measures for the three features are listed below:

1. The $(n + 1)^{th}$ segment is said to be similar in intensity to the n^{th} segment if

$$I_n - t \leq I_{n+1} \leq I_n + t$$

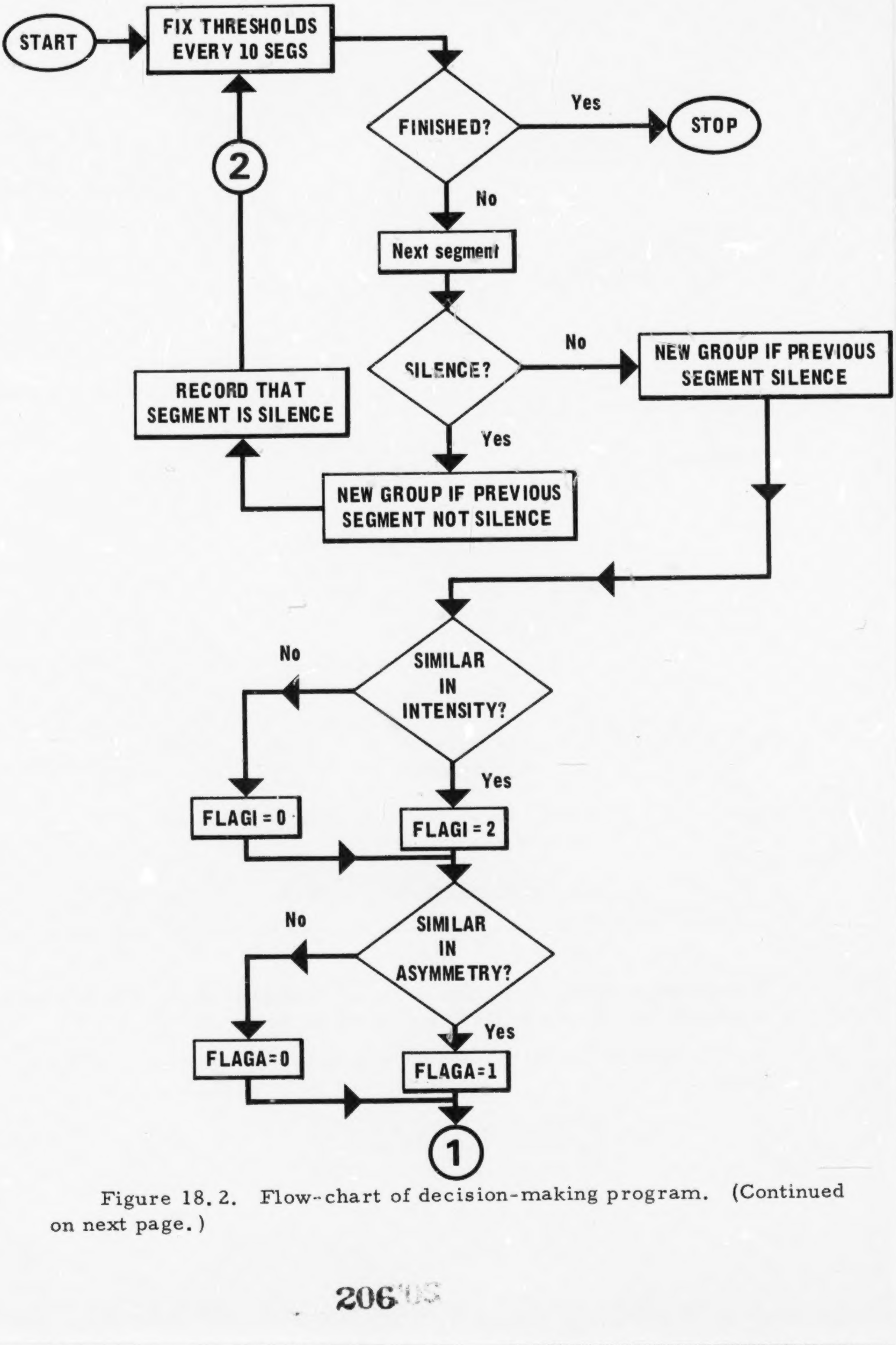

START
FIX THRESHOLDS EVERY 10 SEGS
FINISHED?
Yes
STOP
No
Next segment
2
SILENCE?
No
NEW GROUP IF PREVIOUS SEGMENT SILENCE
RECORD THAT SEGMENT IS SILENCE
Yes
NEW GROUP IF PREVIOUS SEGMENT NOT SILENCE
SIMILAR IN INTENSITY?
No
Yes
FLAGI = 0
FLAGI = 2
SIMILAR IN ASYMMETRY?
No
Yes
FLAGA=0
FLAGA=1
1

Figure 18.2. Flow-chart of decision-making program. (Continued on next page.)

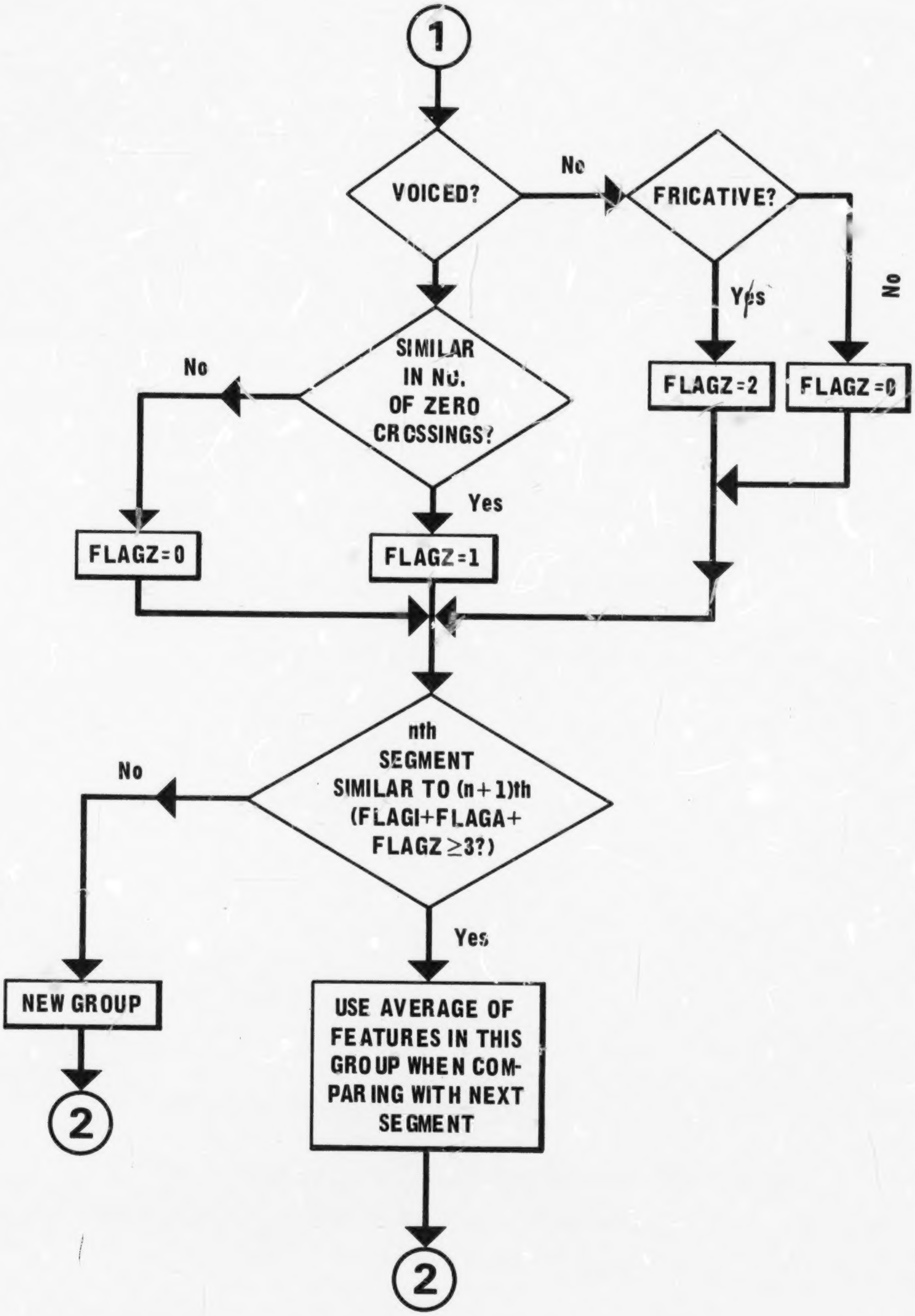

1
VOICED?
No
FRICATIVE?
Yes
No
SIMILAR IN NO. OF ZERO CROSSINGS?
No
Yes
FLAGZ=2
FLAGZ=0
FLAGZ=0
FLAGZ=1
nth SEGMENT SIMILAR TO (n+1)th (FLAGI+FLAGA+FLAGZ ≥3?)
No
Yes
NEW GROUP
USE AVERAGE OF FEATURES IN THIS GROUP WHEN COMPARING WITH NEXT SEGMENT
2
2

Figure 18.2. (Continued from preceding page.) Flow-chart of decision-making program.

or $I_n - t \leq I_{n+2} \leq I_n + t$

where $t = I_n / 8$ if $I_n / 8 \geq \frac{I_{max}}{32}$

and $t = \frac{I_{max}}{32}$ if $I_n / 8 < \frac{I_{max}}{32}$

I_{max} being the max intensity in 500 msec.

2. The $(n + 1)^{th}$ segment is said to be similar in waveform asymmetry to the n^{th} segment if

$A_n \varepsilon P$ and $A_{n+1} \varepsilon P$ where $P = \{ A \mid A > W \}$

or $A_n \varepsilon S$ and $A_{n+1} \varepsilon S$ where $S = \{ A \mid -W \leq A \leq W \}$

or $A_n \varepsilon N$ and $A_{n+1} \varepsilon N$ where $N = \{ A \mid A < -W \}$

or if $A_n \varepsilon P$ and $A_{n+1} \varepsilon S$ and $A_{n+2} \varepsilon P$

or if $A_n \varepsilon N$ and $A_{n+1} \varepsilon S$ and $A_{n+2} \varepsilon N$

3. The $(n + 1)^{th}$ segment is said to be similar in number of zero crossings to the n^{th} segment if

$$Z_n - \frac{Z_n}{4} \leq Z_{n+1} \leq Z_n + \frac{Z_n}{4}$$

or

$$Z_n - \frac{Z_n}{4} \leq Z_{n+2} \leq Z_n + \frac{Z_n}{4}$$

where $Z_n / 4$ is replaced by 1 if $Z_n / 4 < 1$

Two segments are grouped as silence or pause

if $I_n \leq$ SILENCE (Table 18.1) and $I_{n+1} \leq$ SILENCE and $Z_n < 10$ and $Z_{n+1} < 10$.

A segment is voiced if

(i) intensity I_n > V/UV threshold (Table 18.1)

or (ii) asymmetry $A_n \varepsilon P \cup N$

Two 12.5 msec. segments are considered fricative if they are unvoiced, and if the number of zero crossings in one $Z_n \geq 45$ while in the other $Z_n \geq 30$.

After the initial decision-making process, the various groups are searched for redundant segments and the final decisions, as to which segments are actually to be deleted, are made. Silence intervals are treated differently from the rest and the provision of treating unvoiced sounds separately from the voiced sounds could also be incorporated. The rate of compression can be changed by varying the parameters of the program.

Removal of Transients and Output

Once the redundant segments to be deleted are known, the problem is how to join the remaining segments together so that there are no undesirable transients. In the case of voiced sounds, the transients can occur in the form of two glottal pulses lying too close or too far apart compared with the regular pitch period resulting in a noticeable irregularity in the speech sound. The locations of the positive maxima, stored during sampling, are in most cases a fairly good approximation to the locations of the glottal pulses. The rule used for transient removal is that if segments n to m (n<m) inclusive are to be deleted, the first zero or negative sample after the positive maximum in segment n-1 is joined with its counterpart in segment m+1.

Three processes proceed simultaneously during this last phase of speech compression. They are: (1) reading the stored speech samples from disc into core, (2) removal of transients, and (3) D/A conversion of the samples of compressed speech at 4 kHz. The output of the D/A converter is zero order held and a low pass filter with a cut-off frequency of 2 kHz is used to smooth the speech waveform. This speech is recorded on an ordinary tape recorder and played back at twice the speed to restore the frequency spectrum.

The lower rate of 4 kHz of D/A conversion has been necessitated by the limit of data transfer rate from disc to core and is strictly a hardware limitation.

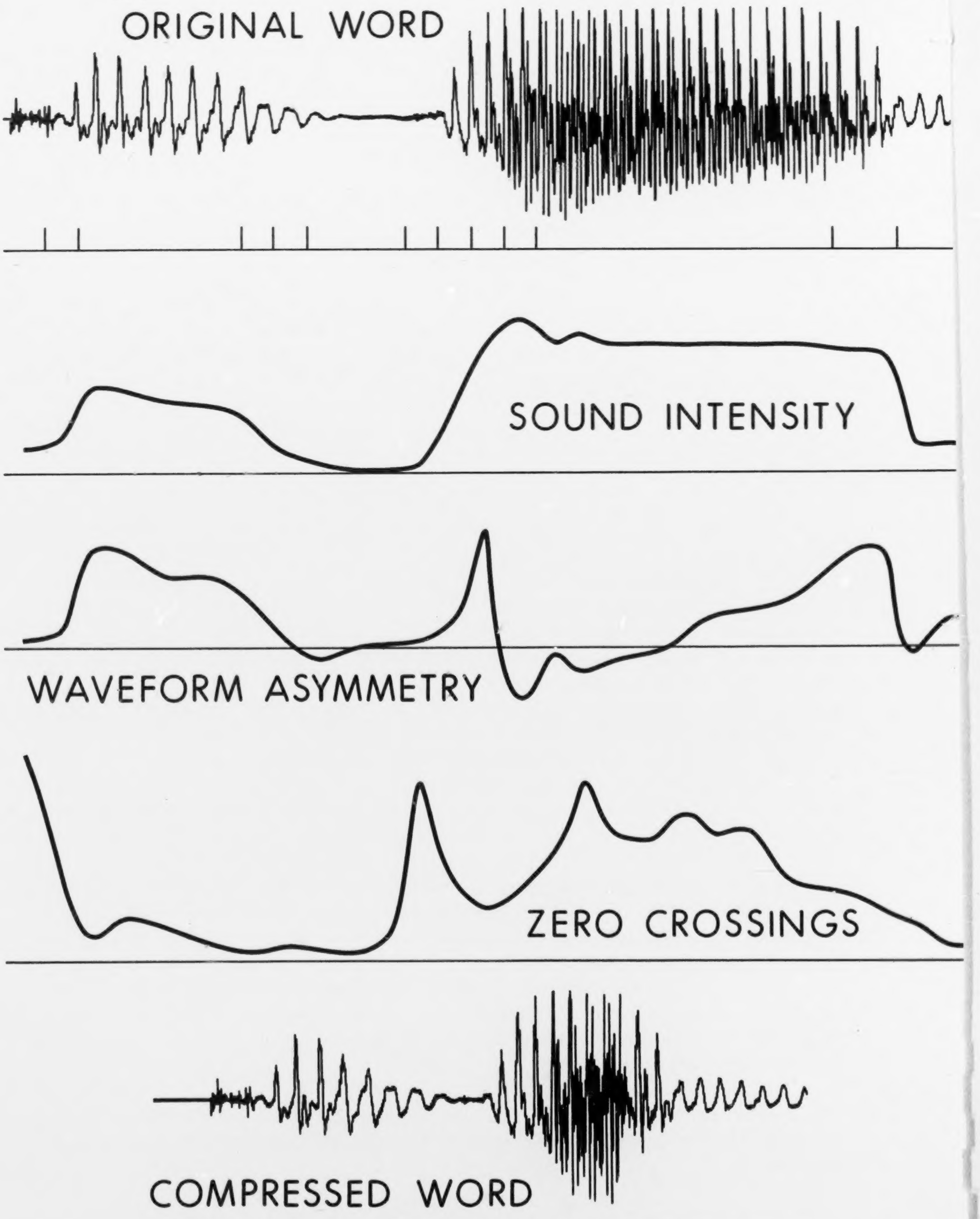


Figure 18.3. Original word, extracted features, and compressed word.

Conclusions

A selective method of compressing speech has been developed in which the total computer time is the time taken for A/D and D/A conversion plus only one-fourth the original duration of the speech for decision making, etc. With the large capacity (500,000 words) high speed discs now available, and remote switch for the tape recorder, the method becomes attractive for compression of continuous spoken passages. It can also be used as a research tool for evaluating the effect of various factors on the intelligibility of compressed speech, and for arriving at a selective rule for compression with optimum intelligibility for a particular word per minute rate of speech.

Fairbanks' sampling method of compression was simulated on the computer and compared with the method reported here. Informal listening tests showed that selectively compressed speech was clearer and free of the low frequency rumble present in Fairbanks' method.

An interesting observation is that the method appears to be independent of the language spoken.

It is also contended that with little modification the same approach can be used for time expansion of speech.

CHAPTER XIX

THE BRAIDED-SPEECH METHOD OF TIME COMPRESSING SPEECH

H. Leslie Cramer and Robert P. Talambiras*

The braided-speech method described here is a refinement of the periodic sampling method for time compressing speech which has been covered elsewhere (Cramer, 1970). Two of the problems inherent in the periodic sampling and discard process which interfere with intelligibility are greatly reduced by the braided-speech approach. This method, however, does not apply to other speech time compression methods not involving time sampling, such as the playback of a recording at a speed greater than that at which it was recorded, and the frequency sampling and division method of the harmonic compressor.

I will quickly review the strategy used for time compressing speech by computer.** A segment of speech may be visualized as a series of time segments of speech records on a piece of tape, as shown in Figure 19.1a. Ordinary computer time compression would simply take the speech record shown in Figure 19.1a and sample it to produce the record shown in Figure 19.1b. This would represent 50% time compression.

The first of the two problems ameliorated by the braided-speech method deals with the desirability of decreasing the sampling length to that of a pitch period of the voice.*** When this is done with the Fairbanks compressor, which is not pitch synchronous, we may find that half of

---

*H. Leslie Cramer, Ed.D., is currently acting director of research with the Peace Corps, 806 Connecticut Avenue, N. W., Washington, D. C. 20525. Mr. Robert P. Talambiras is Vice President, Adage, Inc., 1079 Commonwealth Avenue, Boston, Massachusetts 02215.

**Scott (1965) and Gerber (1968) have both described this in greater detail.

***This is covered in greater detail in Cramer (1968).

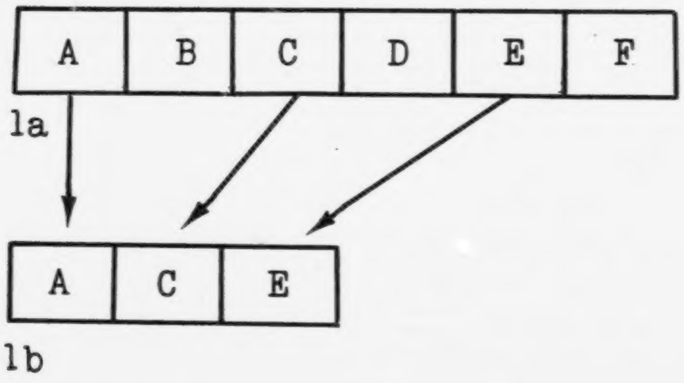


Figure 19. 1. la shows a schematic diagram of a series of time segments of a speech record on tape. lb represents samples taken from speech record shown in la and abutted. This represents 50% time compression.

the end of one pitch period is abutted to the beginning half of the next period within one sample block. Thus there are two half pitch periods to perceive, a quite different thing from that intended. The effect is to raise the apparent pitch of the speaker, just as we would do in playing a record at a fast speed. Fairbanks and Kodman (1957) observed this phenomenon and recommended that in order to prevent a pitch rise and the possibility of the sampling frequency obtruding on the first formant of the speech being processed, a sampling interval of 20 to 40 milliseconds (msec.) be generally used. With a compression ratio above 50%, or twice normal, we are then discarding more than we sample. In a discard over 40 msec. long, we lose over four pitch periods of the speaker's voice. If the speaker happened to be in the middle of a transition between diphthongs or in a glide, a serious and very noticeable distortion would be introduced by leaving out such a long portion. It is this process that sometimes leads to startling transformations of one phoneme into another. In a moment we will return to this problem, but first I want to deal with the second problem, that of bringing two samples, such as A and C in Figure 19. 1, together without causing a lot of noise. Scott (1965) blended a small portion of one sample with the beginning of the next sample to minimize discontinuities. Cramer (1968) reports that at short discard intervals, however, a background tone caused by the frequency of interruption is clearly discernable.

It occurred to us sometime ago when we began computer time compression of speech that, rather than blending just the edges of the sample, that is, the last nine digitized samples of one block with the first nine of the beginning of the next block, as Scott (1965) did, why not blend one complete time block sample into another? This approach led to what we call "braided-speech."

It may be realized that by smoother or more complete blending of one sample into another, short samples, only one pitch period long, may be taken from speech glides. This will allow the time-compressed glide to still glide, as every other pitch period is there at 50% ratios rather than having a large block with several pitch periods missing. A further benefit is that plosives and stops, which are often of both short duration and low energy level, have a higher probability of being sampled sufficiently for good perception.

Almost everyone concerned with producing high quality time-compressed speech has mentioned in one way or another the problem of abutting successive speech samples. Gabor (1947) used graded filters to smooth the edges of samples of motion picture film track recordings being scanned by microscope lenses in a very early speech compressor. Fairbanks, Everitt, and Jaeger (1954, 1959) lifted the tape from one head as the tape was being brought in contact with the next successive head to achieve blending. Springer (1961a) proposed having diagonal gaps scan a tape, rather than vertical gaps perpendicular to the tape, to facilitate blending each sample into the next. This is supposed to render a blending on pickup similar to that which takes place in the portions of speech on each side of a break in a tape when a diagonal splice in a tape goes over a playback head. Scott (1965) blended the last nine digitized samples of one block of speech with the first nine samples of the next block by taking 90% of the ninth record from the end of the first block of samples and blending it with 10% of the first sample of the next block of samples. Likewise, 80% of the eighth to last sample was combined with 20% of the second sample in the next record, etc. Graham (1970) reported on a sophisticated photocell and flip-flop arrangement to gradually blend one record into the next, which works very well.

The problem clearly has been one of getting the edges together smoothly without introducing noise by either cutting the speech waveform at a peak level, or introducing the frequency of interruption. This seems to be one of the most critical problems in terms of making the speech easily understood. Those of you who have heard relatively unblended speech know that it has a roughness or abruptness to it. As it hits the ear, it drives the hearing threshold to a higher level so that the ear is not as responsive. Then by the time the threshold is lowered a little so that the ear can hear the next sound, the ear gets "whomped" again by another sharp discontinuity. Voor (1962) assumed that some of the sharp discontinuities were accidents, unfortunate but impossible to eliminate, inherent in the sampling process. He dealt at some length in his master's thesis with the need for low impedance headphones to eliminate the "ringing" caused by the steep wave fronts occasionally produced when sampling cut the speech waveform at a peak value.

In order to impart the speech processing subtleties possible with the braided-speech method, it will be necessary to give a brief explanation of the computer processing that takes place. There are several steps that occur. First is the conversion of the analog output of a tape recorder to a digitized record. Then follows a series of operations on the digitized speech record within a computer, followed by reconversion of the desired digital record to analog output and rerecording the new braided-speech record.

We will return now to the first step, getting the speech record, which is to be compressed, into the computer. The original speech passages recorded at 15 ips were played on a Tandberg stereo tape deck at half speed (7 1/2 ips) into a Krohn-hite filter. This was a low pass filter with 24 db per octave attenuation and was set to 4,500 Hz. This effectively filtered the speech at 9,000 Hz to prevent beating with the 12 kHz sampling rate of the computer. The computer sampling rate of 12 kHz real time, applied to the speech played in at half speed, gave us an effective 24 kHz sampling rate of the speech. Each sample was encoded into 12 bits plus sign giving $\pm 2,048$ levels and therefore a 66 db dynamic range. In order to take full advantage of the 66 db dynamic range and the low signal-to-noise ratio, a small amount of peak clipping was performed by setting the input level high enough to occasionally overdrive the ± 40 volt input range of the computer to as high as ± 70 volts on peaks which were then encoded at the maximum value allowable of 40 volts. This treatment gave the effect of preemphasis on consonants, particularly the fricatives, by giving them a sharpness and clarity otherwise lacking.

From the filter, the signal went directly into the analog-to-digital converter within the Ambilog 200 computer.* The speech input was then digitized at 12 kHz with each sample accurate to 12 bits. The entire speech record is digitized and written in gapped IBM-compatible magnetic tape format at a writing density of 556 bits per inch before any subsequent processing takes place.

In processing the digitized speech samples in the computer to produce braided-speech, we start with the digitized speech record as represented in Figure 19.2. The speech is blocked or grouped in 10 msec. long temporal blocks. Each of these 10 msec. sample blocks is composed of 240 digitized samples encoded at 12 bits. This roughly is the temporal length of one speech pitch period of the reader of our passages. We therefore can conceive of a series of pitch periods, one in each block. The first thing that is done in the computer is to successively multiply each of the 240 digitized samples in block sample A by a series of fractions decreasing in size. The first sample is multiplied by 240/240, the second by 239/240, the third by 238/240--to the 240th sample by 1/240. Essentially what we're doing is changing whatever speech is in sample A from a full scale level to zero volume or amplitude level. We have sample A coming in and sweeping from maximum value to zero in a linear fashion. In block sample B, the reverse is done: each of the 240 digitized samples is multiplied by the factors in reverse. That is, the first sample in B is multiplied by 1/240, the second by 2/240, the third by 3/240, etc., right on up to the 240th sample by 240/240 constituting 100% amplitude level. Each successive pair of blocks (C-D, E-F, etc.) are similarly treated. This processing does what one could do, if he were fast enough, with a linear volume control, turning it down smoothly during block A, then turning the volume smoothly back up to maximum through block B, then down again during block C, etc., alternating through each successive block of 240 samples.

---

*The Ambilog 200 is a hybrid computer produced by Adage, Inc., 1079 Commonwealth Avenue, Boston Massachusetts. It is described in Grandine and Hagan (1965).

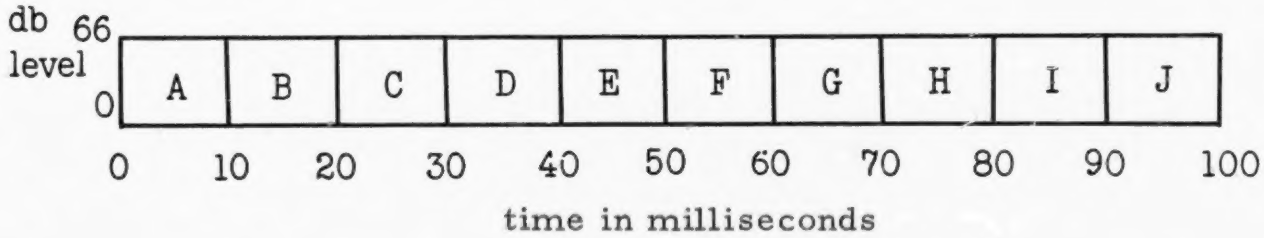

db level
66
0
A
B
C
D
E
F
G
H
I
J
0
10
20
30
40
50
60
70
80
90
100
time in milliseconds

Figure 19.2. Schematic representation of a speech record digitized and blocked into 10 msec. intervals for processing in digital form. Each block represents 240 samples digitized into 12 bit quantities + sign.

Figure 19.3 shows this first step in the braiding process in a schematic diagram.

The next process is to algebraically add sample A and B together so that sample A is going from full level down to zero while sample B is going from zero up to maximum. We bring sample C over next to B and blend sample D onto it. Sample C is going from maximum down to zero while sample D is going from zero up to maximum. This scheme, shown in Figure 19.4, would represent 50% compression. We actually have 100% of the speech present in a 50% compression.

Similar processing can also be done in the dichotic domain where in channel B, sample D is blended onto A, and H onto E, while in channel A we likewise have samples F blended onto C and J onto G. See Figure 19.5

At four times normal or 75% compression ratio, with a dichotic braiding of the speech, we have 100% of the speech present except for the fact that some of it is at zero level. Overlaying the pitch periods or time interval blocks on top of one another, especially in a dichotic presentation as shown in Figure 19.5, and following the pitch periods through a period of time, creates a pattern resembling a four-strand braid, hence the name "braided-speech."

It may be noted in Figure 19.5 that channel B is offset by 120 samples from channel A. This is done to insure that the junction between digitized blocks D and E, which abut each other, are temporally spaced between the junction of blocks B-C and F-G in the adjacent channel. This amount of offset represents 5 msec. of time. Both zero offset and 10 msec. offset were tried in addition to the 5 msec. offset with no discernible difference in effect, and so the 5 msec. offset was used for processing final tapes presented to listeners as it seemed to be the most logical approach.

One other factor should be noted at this point. Gabor (1947) experimented with linear, sinusoidal, and various other blending approaches. He maintained that a normal probability curve superimposed on each sample to be combined was the smoothest. The first computer compressions which we tried used this schedule but we noticed a very pronounced modulation effect. On examining the logic of this approach, it was evident that the summation of each of the two fractional parts representing each sample level under the curve did not add up to 1. At this point a linear approach was tried and the modulation effect was no longer noticeable.

The last step in the computer processing was to write one of the dichotic channels in every other gap on a magnetic tape, then rewind the tape and write the second channel in the gaps skipped when recording the first

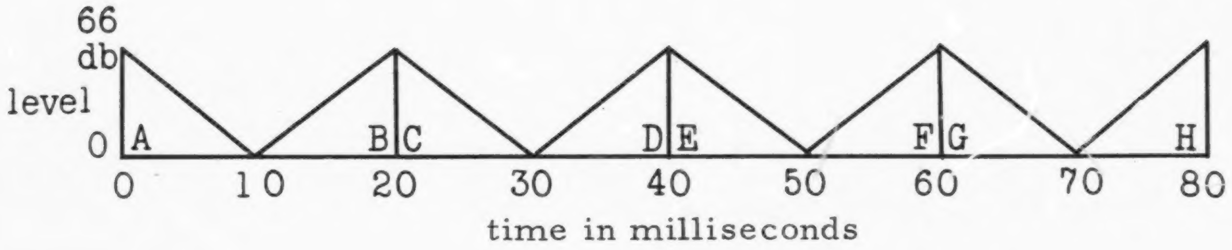


Figure 19.3. Schematic diagram of first step in process of braiding speech. Each lettered segment represents a pitch period of speech multiplied by a series of decreasing and increasing fractions changing in direction of magnitude with each 240 digitized samples or 10 msec. of speech. This represents procedure for 50% compression.

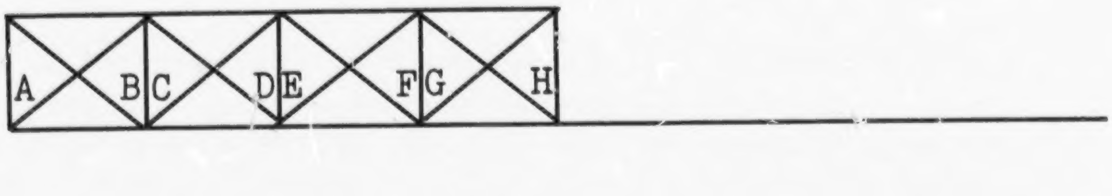
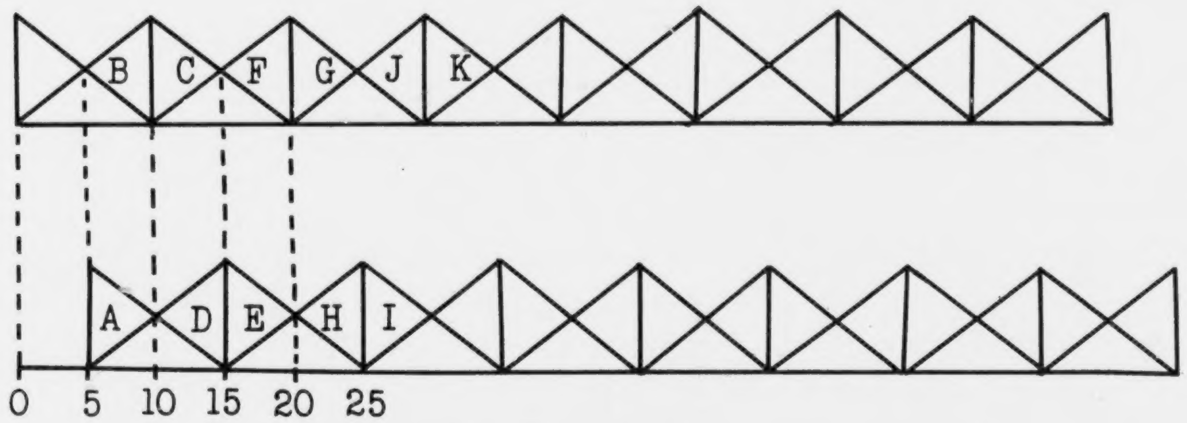


Figure 19.4. Schematic diagram of second step in process of braiding speech. Every other block of speech in Figure 19.3 is brought over onto the previous block; B onto A, D onto C, etc., and "braided" by algebraically adding the two analog signals in each of the two blocks so combined. The straight top line represents the sum of each of the pairs of blocks added.



time in milliseconds

Figure 19.5. Schematic diagram of process of dichotic braiding of speech. The first step is the same as shown in Figure 19.3. The second step shown here represents 80% compression or four times normal rate. The 5 msec. offset between channels is explained in the text.

channel. Following this the magnetic tape was played out through the Ambilog unit which controlled the reading of two separate records from the tape and provided the specified amount of offset between the channels. The two channels were fed from the Ambilog unit into separate digital-to-analog (D to A) converters, and from these through two Krohn-hite interpolation filters. On output, since two channels were being played at once, the frequencies which were fed into the computer at half speed were further divided by a factor of two so that the speech output was in one-quarter real time. It was therefore filtered at 2, 250 low pass and 75 Hz high pass to smooth or interpolate between steps in the output from the D to A converters and to eliminate beating with the fundamental frequency of the voice being processed. From the filters the analog signal was fed directly to the Tandberg stereo tape recorder and recorded at 1 7/8 ips. For normal playback, it was then played at 7 1/2 ips to restore the voice to normal frequency in real time. See Figure 19.6 for a schematic diagram of the input, processing, and output equipment configuration.

In testing braided-speech by comparing it with speech compressed on Fairbanks' compressor, three different types of presentation were tried. One of these was monotic, that is, a single channel of the dichotic speech was presented to both the left and right ears of a group of listeners. The second was dichotic presentation, that is, one channel to the left ear and the other to the right ear. The third method was accomplished by mixing the left and right channels algebraically to derive a single center channel output which was then fed to both the left and right ears (see Figure 19. 7).

The maximum intelligibility was discovered to be about the same for all four presentation treatments at 50% compression ratio, but as the compression ratio was increased above 50%, the dichotic version was more intelligible than feeding separate dichotic channels to each ear. This is contrary to an earlier report, namely that of Gerber and Scott (1970) We cannot explain why; at this point we can only report our findings. * The differences in findings in the two studies may represent differences in the particular length of sample or the amount of the offset of samples from one channel to the other, or possibly are due to differences caused by braiding rather than blending.

---

*Full details of the experimental design, subjects, procedures, and data analysis are expected to be forthcoming in the Journal of Communication.

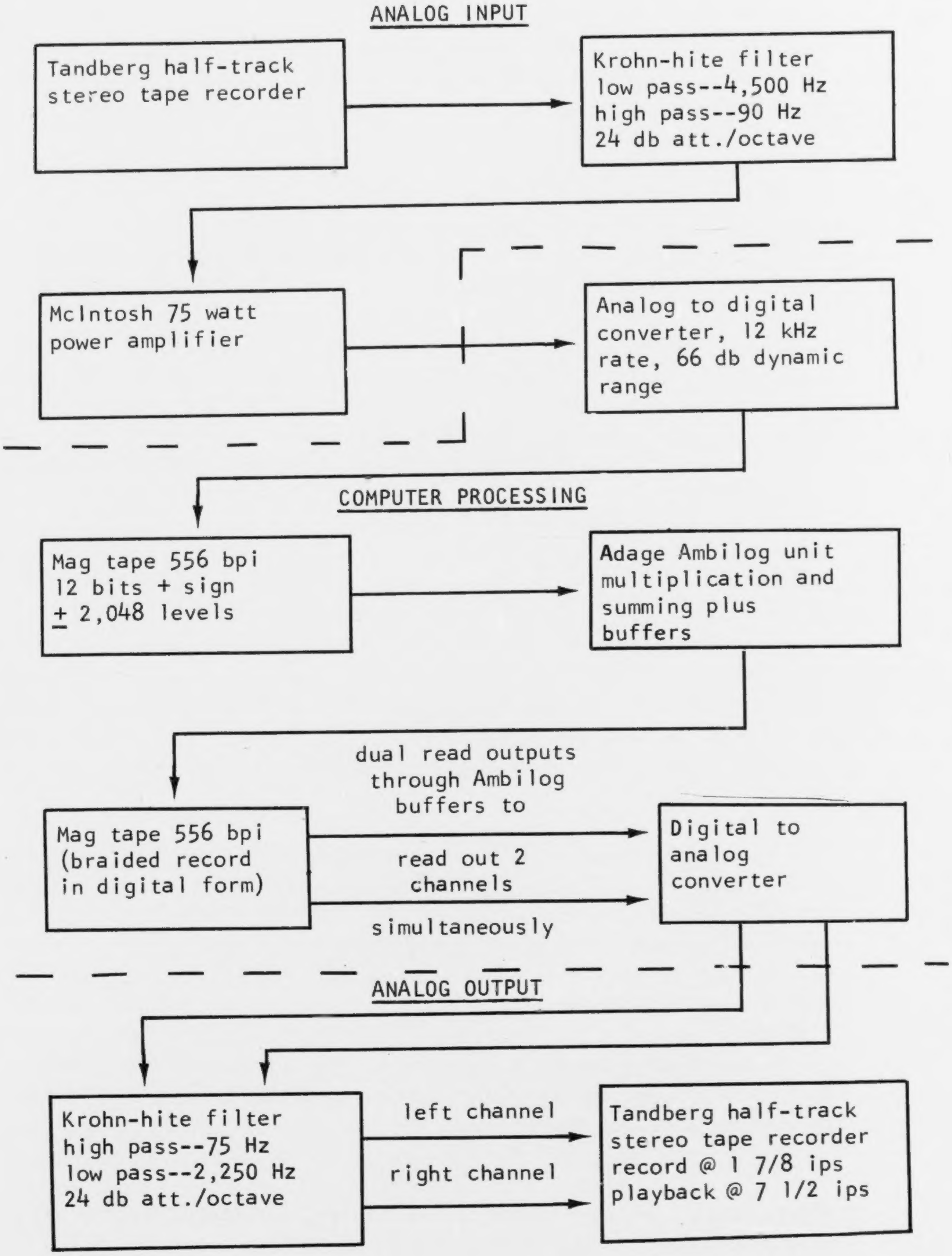

ANALOG INPUT
Tandberg half-track stereo tape recorder
Krohn-hite filter low pass--4,500 Hz high pass--90 Hz 24 db att./octave
McIntosh 75 watt power amplifier
Analog to digital converter, 12 kHz rate, 66 db dynamic range
COMPUTER PROCESSING
Mag tape 556 bpi 12 bits + sign ± 2,048 levels
Adage Ambilog unit multiplication and summing plus buffers
dual read outputs through Ambilog buffers to
Mag tape 556 bpi (braided record in digital form)
read out 2 channels
simultaneously
Digital to analog converter
ANALOG OUTPUT
Krohn-hite filter high pass--75 Hz low pass--2,250 Hz 24 db att./octave
left channel
right channel
Tandberg half-track stereo tape recorder record @ 1 7/8 ips playback @ 7 1/2 ips

Figure 19.6. Block diagram of input, throughput, and output equipment used in producing time-compressed speech with the Ambilog 200 Computer by the braided-speech method.

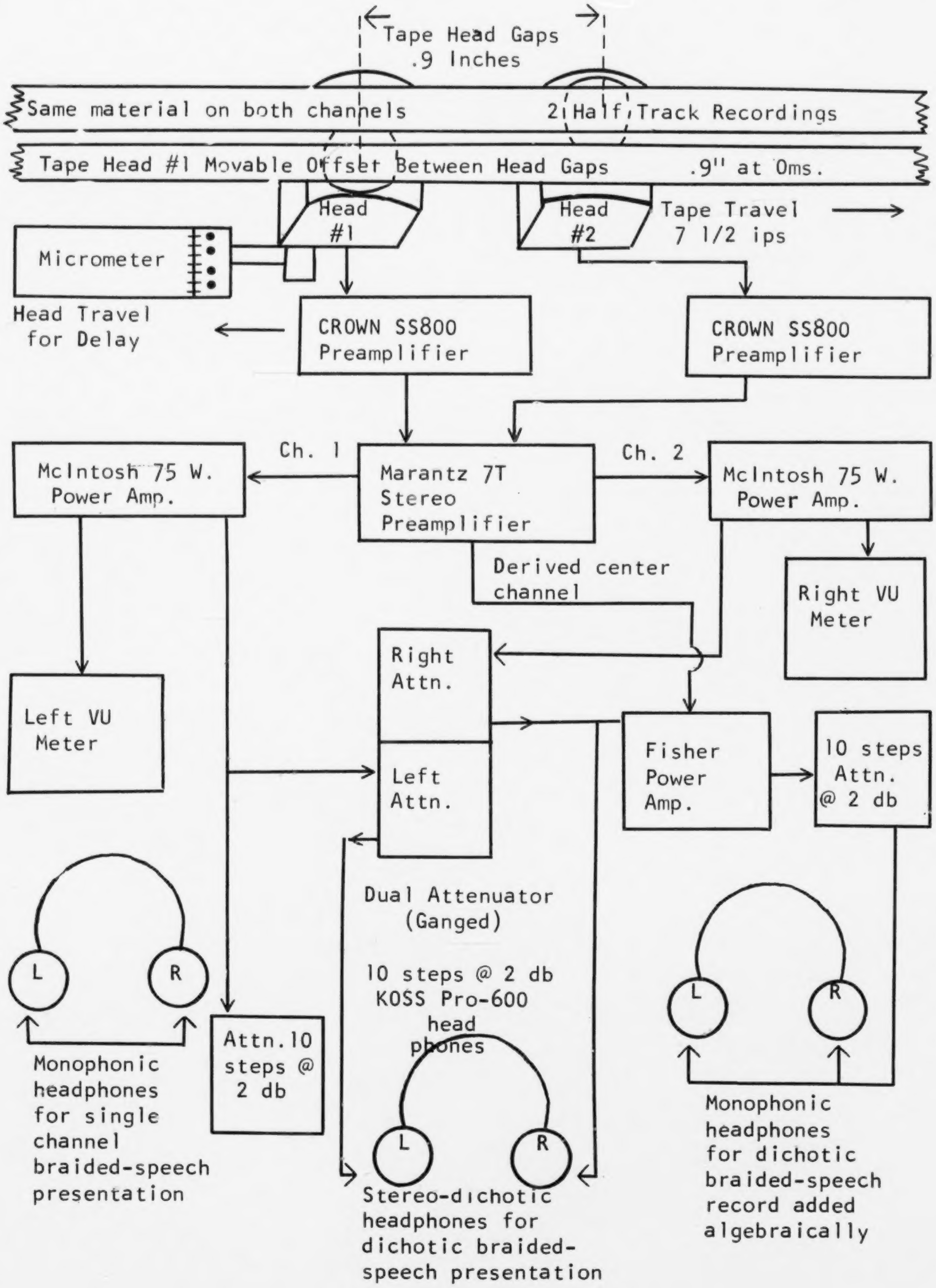

Tape Head Gaps
.9 Inches
Same material on both channels
2 Half Track Recordings
Tape Head #1 Movable Offset Between Head Gaps .9" at 0ms.
Head #1
Head #2
Tape Travel
7 1/2 ips
Micrometer
Head Travel for Delay
CROWN SS800 Preamplifier
CROWN SS800 Preamplifier
Ch. 1
Ch. 2
McIntosh 75 W. Power Amp.
Marantz 7T Stereo Preamplifier
McIntosh 75 W. Power Amp.
Derived center channel
Right VU Meter
Right Attn.
Left VU Meter
Left Attn.
Fisher Power Amp.
10 steps Attn. @ 2 db
Dual Attenuator (Ganged)
10 steps @ 2 db
KOSS Pro-600 head phones
L
R
L
R
L
R
Attn.10 steps @ 2 db
Monophonic headphones for single channel braided-speech presentation
Stereo-dichotic headphones for dichotic braided-speech presentation
Monophonic headphones for dichotic braided-speech record added algebraically

Figure 19.7. Equipment configuration for 3 different types of braided-speech presentation: (1) monotic, (2) dichotic, and (3) derived center channel.

ACKNOWLEDGMENT

The research and development reported herein was performed pursuant to a contract (OE 5-10-239) with the United States Department of Health, Education, and Welfare, Office of Education, under the provisions of the Cooperative Research Program, as a Project of the Harvard University Center for Research and Development on Educational Differences. Reproduction in whole or in part permitted for any purpose by the United States Government.

CHAPTER XX

THE MEASUREMENT OF LISTENING COMPREHENSION

David B. Orr*

Introduction to the Sessions of the Day**

It occurs to me that some of the material that we heard yesterday had somewhat conflicting results in it. I know, for example, that the work Herb Friedman and I did at the American Institutes for Research showed a very clear practice effect on the comprehension of high speeded speech, whereas Dr. Woodcock reported that he had found no such effect. There were also some other negative results reported, and I am all in favor of presenting such negative results because I feel that these will help us to define this field and find out where we are.

It occurs to me, however, that much of the problem of conflicting results may revolve around the problem of measurement, and I want to enter a plea with those of you who are working in this area to do some creative thinking about the problem of how you measure comprehension in these various experiments. I have the feeling that some of our contradictory results are arising from this particular problem. In a sense we are measuring with something like "rubber yardsticks."

I would like to offer several observations along these lines, and I don't intend to suggest that I have thought of things that nobody else has thought of here, but I think perhaps taking the opportunity to review briefly the shortcomings of some of our measurement approaches is appropriate.

Most typically, the multiple-choice listening analog to the standard reading comprehension test is the thing that gets used to determine comprehension. There are several particular problems with this approach. In the first place, it's very difficult to produce a research instrument which is a psychometrically adequate test. We need really to go through the same kind of procedure that the test publisher goes through in order to come up with an instrument which has demonstrated reliability, validity, practicality, good-item forms, and a high-average intra-item correlation, and all those other sorts of things with which a psychometrician is concerned. It doesn't suffice to pooh-pooh and say: "Well, this is just a research instrument, and we don't have to do all those things." We're just as concerned with reliability and validity of measurement as anyone else. So there are these kinds of problems with multiple-choice tests--and, of course, with almost all measurement approaches.

---

*Dr. David B. Orr is President and Treasurer of Scientific Educational Systems, Inc., 910 17th Street, N. W., Washington, D.C. 20006.

**These remarks were delivered extemporaneously by Dr. Orr as the introduction to the sessions of the day.

There is also a problem which can be called the "domain" problem. Ideally, the test ought to be an unbiased and representative sample from the domain of material which was presented to the listener for comprehension. If it isn't an unbiased sample of that domain, then whatever way we cumulate the examinee's responses gives us a biased representation of his comprehension. And, it can prove exceedingly difficult, as those of you who have tried will testify, to write a set of questions which is really an unbiased and representative sample of the information contained in a passage of material. But nonetheless, we have to try; if we are going to use the multiple-choice approach, this is important.

A third point with respect to multiple-choice tests has to do with prior knowledge. Questions which the individual can answer on the basis of knowledge that he got elsewhere than listening to the passage are irrelevant with respect to measuring comprehension, at least of that particular presentation. It's extremely difficult to rule out the effects of prior knowledge. I have the feeling, however, that in many cases we don't try as hard as we might. Herb Friedman and I, in our work, used a little known high school text in precolonial English history (The English People of Precolonial Times, 1500-1600, or something like that). I have to admit that the material in there was novel to me anyway. It's worth trying.

Another method of measuring comprehension which has recently come into prominence is the cloze test. The cloze test--most of you are probably familiar with this by now--is a domain test; there's not much argument that it's a domain test because the passage is taken and every fifth word (or every nth word) is deleted, and the subject is asked to fill in the blanks. Clearly it is a domain test because here's the passage being presented with these deletions, but it bothers me a little bit somehow. I don't think it obviates the problem of prior knowledge--but the prior knowledge is of a different sort here. The prior knowledge is associated with not only the prior knowledge of the content of the passage, but it also is dependent to some degree on the structure of language which the individual has built up over a period of time. I'll be perfectly frank with you, I'm no linguist. I don't know very much about the probabilities of a given word following four other words, except that there seems to be a lot of predetermination in this--if you say ham, most people will say eggs--and I simply don't know how this sort of thing interacts with the measurement problem. All right . . .

> (From floor: I hate to interrupt in the middle of this but something I think a lot of people have been overlooking is that old technique of just finding empirical levels for our tests. I notice the Nelson-Denny Test has been reported at statistical chance levels, but people are forgetting--why not just take control groups and get your empirical level from the samples before you run these tests, because I have found in using the Nelson-Denny that the empirical level is above the chance level, due to prior knowledge. This is an example and it might be well for some of the others to try to do this.)

I agree with this suggestion although there are obvious difficulties in implementing it. As a matter of fact, I think I have mentioned what I think are the key problems. Now, I want to offer a few suggestions--lay them on the table, quivering sort of like, for you to dissect if you wish on your own time, and that was going to be one of the suggestions that I was going to make. I think that there is probably little excuse for using any of these techniques without some kind of correction for chance and without some kind of an estimate of what prior knowledge is on an empirical basis (such as a prior administration of a test). I want also to suggest a couple of approaches to measurement that might prove fruitful although they don't necessarily eliminate all of the problems. It seems to me that part of our problem is in defining comprehension, and if you are willing to take a limited definition of comprehension you can talk about some things like the following: let's suppose you develop a response device--let's suppose that this response device looks something like Figure 20. 1.

We draw on the blackboard a box, if you will, with four switches on it. Let's suppose you also develop a set of cards which depict pictorially various situations and which simply slips down on top of the box with **holes** for the switches. Let's suppose that you offer then, as your comprehension material, a set of instructions, a set of relationships in which the individual is required to integrate the information that is being supplied to him in the passage with the picture on the box and indicate a response by throwing a switch or a combination of switches. There are enough combinations and enough possibilities in this kind of approach to provide all the difficulty which you would possible want. At the same time, the answers to this, besides giving you electromechanical recording of the response time which I think is also interesting, can give you a

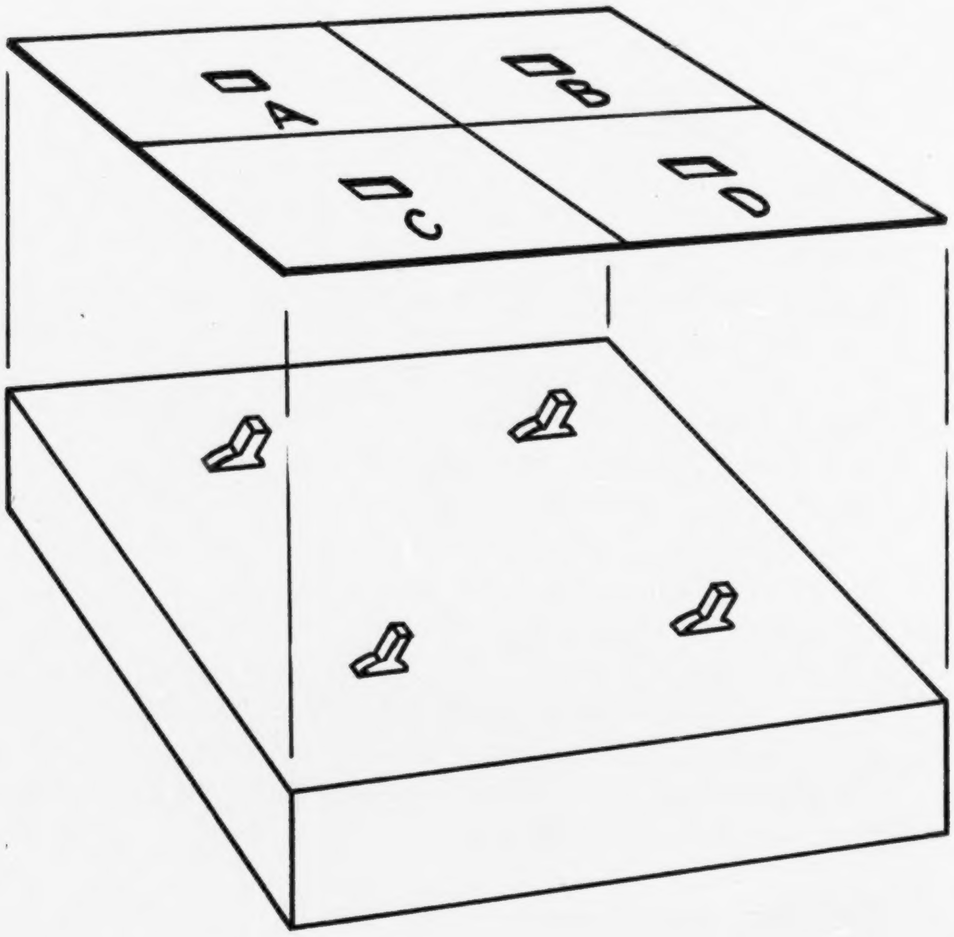


Figure 20.1. Response device.

response which doesn't depend upon prior knowledge as much as--you can never eliminate it entirely, of course--the usual passage comprehension approach. O. K. That's something to think about.

The other thing that I'd like to suggest for you to think about is what we might call the "match paraphrases" approach to comprehension. This can be done in either of two fashions. You can either talk about a model, which the individual learns and which consists of maybe a couple of sentences (it's not a very complicated model), and you then present auditorially paraphrases of that model and ask the individual to indicate which paraphrase best matches the information in the model. Now, you've got to give him a minute or two of study time on the model. This approach is similar to the psychometric technique of testing memory for sentences, for example, where you give the individual a few seconds of study time for some material and ask him to reproduce it. In this case, what we're saying is that "prior knowledge" is supplied by the model. The comprehension test is "do you understand the paraphrase well enough to match it to the model." The second approach to this, of course, is to present a short series of paragraphs and ask the individual to pick the two paragraphs which said essentially the same thing.

Now, these approaches are perhaps loaded with memory to some extent. Nonetheless, I think approaches of this kind, or other creative approaches, to the measurement problem represent responses to a crying need in our efforts to try to measure the impact of rate-controlled speech and of listening to auditory presentations in general. With that, I'll stop taking up your time, and get on to introducing the first speaker of this session.

CHAPTER XXI

EFFECTS OF MOTIVATION AND WORD RATE ON AURAL COMPREHENSION

Carson Y. Nolan and June E. Morris*

The purpose of this study was to test the effects of varying motivation on comprehension of material presented at three different word rates. A secondary finding of earlier research (Nolan, 1968), was that, when heard under motivated conditions, material of three different types was comprehended significantly better when heard at a normal rate than when heard at a rate slightly faster than normal.

Numerous studies (Foulke, 1968a; Foulke & Sticht, 1967a; Henry, 1966; Jester, 1966; Loper, 1966; Wood, 1965) have shown that the relationship between comprehension of aural material and word rate is a negative one. However, the point at which the decrease in comprehension becomes significant has never been clearly identified and has been reported at various rates. Generally, it has been thought that comprehension was not significantly affected at rates below 250-275 words per minute (wpm). Because of this, Nolan (1968) did not anticipate finding significant rate differences in his study in which only moderately compressed material, 225 wpm, and normal material, approximately 175 wpm, were used. When significant rate differences favoring the normal rate did occur, the role of motivation immediately became suspect; all Ss in this study having worked under motivating conditions. Nolan suggested that real differences may exist even at low levels of compression which only become apparent when Ss perform optimally as they might be expected to do when working under motivated conditions. The current study was designed to follow through on the earlier study by testing the role of motivation on the learning of aural material heard at normal and compressed rates.

---

*Dr. Carson Y. Nolan is the Director of the Department of Educational Research and June E. Morris is a Research Associate at the American Printing House for the Blind, Louisville, Kentucky 40206.

Schools providing both high school and elementary *S*s were the Nebraska School for the Visually Handicapped, the West Virginia Schools for the Deaf and the Blind, and the Wisconsin School for the Visually Handicapped. Additional elementary *S*s were required and these were obtained at the Texas School for the Blind. With the exception of the Wisconsin school, all regular students enrolled in these schools at the appropriate grade levels who were present at the time the data were collected participated in the study. It was not necessary to use all of the students available at the high school level in Wisconsin so those students used were selected randomly from the total high school population.

Braille and print readers were used as they occurred naturally in the schools; however, initially a stratified-random sampling technique (Guilford, 1956, pp. 158-159) was used in assigning *S*s within a grade level (4-5, 6-7, 8-12) to treatment groups. For each school this was accomplished by first randomly assigning all braille reading students within a level to the treatment groups and then randomly assigning the print reading students within the same level to the treatment groups. Where not all groups within a level contained an equal number of students of the same reading medium from a site, assignment of *S*s of the other reading medium was made in such a way that *S*s within a school were equally or nearly equally distributed among the six treatment groups. By assigning *S*s to groups in this manner, the groups initially contained proportions of braille and print *S*s representative of the school populations from which they were drawn.

Absenteeism and *S*s lost due to their inability to perform the required task in the allotted time resulted in some attrition within the initial groups. Having anticipated some losses, more *S*s were assigned to most of the groups than were required by the experimental design. After all data were collected, surplus *S*s were randomly omitted within grade levels.

Material

Two literary selections were used that had been part of previous research examining the parameters of learning by listening. "A Battle Over the Teacups" (Derleth, 1957) was heard by *S*s at the high school level. The story contained 1,970 words and had a reading difficulty appropriate for eighth- and ninth-grade students, as determined by Flesch's readability formula (Flesch, 1951). This difficulty level is typical of material appearing in high school literature texts. "Notchtail" (Stauffer, Burrows, & Jones, 1962) was used by *S*s at the elementary level. The version used contained 2,114 words, the original version having been edited to shorten it slightly. Reading difficulty for this story, as computed by the Flesch formula, indicated it was appropriate for use by sixth-grade students.

Method

Design

Similar studies were designed for students in grades 4-7 and for students in grades 8-12. In each, Ss listened to literary material presented at one of three rates under either motivated or unmotivated conditions. Factorial designs involving random groups were used.

For the high school group, a 3 x 2 design was used with rate of listening and mode of listening being the treatments. The three rates of listening were 175 wpm (actual), 225 wpm, and 275 wpm. The two modes of listening were either under motivated or unmotivated conditions.

The 3 x 2 x 2 design used in the elementary grades was similar to the high school design but with a grade level factor included. The latter was included because previous research (Nolan, 1966) using the same materials, revealed grade level differences in the amount learned by elementary students. The two levels used in the current study were grades 4-5 and grades 6-7.

Subjects

All Ss were legally blind students enrolled in residential schools for the blind who were assigned to regular classes. One hundred twenty Ss were used at the high school level; 20 being in each of the six treatment groups. One hundred eight Ss were used at the elementary level; 18 being in each of the six treatment groups. Subjects in each of the elementary groups were divided equally so that half were from grades 4-5 and half from grades 6-7. Table 21.1 describes the composition of the groups.

TABLE 21.1

SUBJECTS AT EACH LEVEL IN EACH TREATMENT GROUP

	Unmotivated			Motivated		
Grade Levels	175 wpm	225 wpm	275 wpm	175 wpm	225 wpm	275 wpm
4-5	9	9	9	9	9	9
6-7	9	9	9	9	9	9
8-12	20	20	20	20	20	20

Both selections were recorded on magnetic tape by a professional reader in the recording studios at the American Printing House for the Blind (APH). These were then compressed to three different rates by a time sampling technique at the Center for Rate-Controlled Recordings at the University of Louisville. The specifications for the resulting tapes were that they contain speaking rates of 175 wpm (actual), 225 wpm, and 275 wpm. Preceding each was a short sample taken from "History of Milling" (Buehr, 1959) recorded similarly and compressed to the same rates as the selections which followed.

Tests of comprehension for both selections were reproduced in braille and large type (18-point by APH standards). The tests for the high school and elementary selections were five choice multiple-choice tests containing 70 and 65 questions, respectively. Reliability for the braille and large type editions of the tests ranged from .91 to .95 (Nolan, 1966).

Procedure

Subjects in motivated groups were told they had a chance to win two prizes. First, that each student earning the highest score on the test in his local group would receive a prize and, second, that all students in a local group would receive a prize if that group had the highest average score among similar groups from all the participating schools. In the interschool competition, groups competed only with similar treatment groups. Students were informed that the prizes would be candy.

All students at a school assigned to the same treatment group worked together. Unmotivated groups were scheduled prior to motivated groups. One examiner could work with three groups during a school day. After determining the day for a given series, i.e., unmotivated high school groups, the order in which the three groups within that series were seen was determined by random means.

With each group, Ss were assembled, instructions read, the sample played, the selection played, test instructions read, and the tests administered. Playing the sample offered an opportunity for Ss to become acclimated to the listening environment, the reader, and the rate of presentation. Tests were given without time limits; however, in a few cases the demands of the schedule and/or school day required dismissing a S before he had completed his test. In such cases, where the student had completed 85% or more of his test the score was prorated and used.

The tests were shipped back to the APH where they were scored and/or checked and the prizes shipped. To avoid feelings of unfair treatment, similar prizes to those for the winners in the motivated groups were sent to participants in the unmotivated groups. This procedure was explained to the administrators of the schools involved so that they could pass the word on to their students as they distributed the prizes.

Results

Results of this study fail to substantiate Nolan's earlier finding (1968) that aural material is comprehended better when heard at normal rates than when heard at 225 wpm. Also, the findings do not support the hypothesis that motivation is related to this occurrence.

Reference to Table 21.2 shows that rate of presentation was the only factor significant (.01 level) to learning aural literary material for high school students. A glance at Table 21.3 verifies that as rate of presentation increased, comprehension decreased. Differences in learning between motivated and unmotivated groups were not significant nor was there any significant interaction between motivation and rate.

TABLE 21.2

HIGH SCHOOL ANALYSIS OF VARIANCE SUMMARY

Source of Variation	df	Sums of Squares	Mean Squares	F
Rate (R)	2	2,926.67	1,463.34	11.90**
Mode (M)	1	76.80	76.80	.62
R x M	2	77.40	38.70	.31
Within Cells	114	14,023.10	123.01	
Total	119	17,103.97		

** Significant at the .01 level

TABLE 21.3

MEANS, STANDARD DEVIATIONS, RANGES, AND NUMBER OF SUBJECTS AT THE HIGH SCHOOL LEVEL

	Unmotivated			Motivated		
	175 wpm	225 wpm	275 wpm	175 wpm	225 wpm	275 wpm
Grades 8-12						
Mean	39.4	38.8	29.1	39.9	35.4	27.2
S. D.	13.0	11.6	11.5	10.5	10.9	8.5
Range	16-56	15-57	9-55	21-60	18-64	11-46
N	20	20	20	20	20	20

Differences significant at the .05 level were found for rate and, as expected, grade level for the elementary group. Table 21.4 shows that these were the only significant differences occurring within this group. As with the high school students, the relationship between comprehension and word rate was a negative one. Table 21.5 shows this to be true for all elementary subgroups.

TABLE 21.4

ELEMENTARY ANALYSIS OF VARIATION SUMMARY

Source of Variation	df	Sums of Squares	Mean Squares	F
Rate (R)	2	1,335.36	667.68	4.66*
Grade Level (GL)	1	966.02	966.02	6.75*
Mode (M)	1	60.76	60.76	.42
R x GL	2	6.24	3.12	.02
R x M	2	48.16	24.08	.17
GL x M	1	420.08	420.08	2.93
R x GL x M	2	70.39	35.20	.24
Within Cells	96	13,741.33	143.14	
Total	107	16,648.34		

* Significant at the .05 level

TABLE 21.5

MEANS, STANDARD DEVIATIONS, RANGES, AND NUMBER OF SUBJECTS AT THE ELEMENTARY LEVELS

	Unmotivated			Motivated		
	175 wpm	225 wpm	275 wpm	175 wpm	225 wpm	275 wpm
Grades 4-5						
Mean	38.0	35.8	31.4	34.8	29.4	24.7
S. D.	13.5	11.1	12.1	12.6	13.9	7.2
Range	16-53	12-46	15-52	15-49	9-48	16-36
N	9	9	9	9	9	9
Grades 6-7						
Mean	41.2	39.4	30.7	43.3	39.0	36.3
S. D.	11.4	17.4	7.4	9.7	11.6	12.0
Range	21-59	15-59	20-43	23-56	19-52	16-54
N	9	9	9	9	9	9

Several t tests were run to analyze further the differences in the means for the different rate groups. These revealed that high school students learned significantly more at both 175 wpm and 225 wpm than they did at 275 wpm while elementary students learned significantly more at 175 wpm than at 275 wpm. All of these differences were significant at the .01 level of confidence.

Conclusions

With the exception of the expected significant grade level difference found for elementary students, comprehension of literary material appeared to be affected only by rate of presentation; comprehension and word rate being negatively related. Blind students at both the high school and elementary levels learned more from material heard at the slower rates. Learning appeared to be unrelated to motivation.

In reviewing Nolan's earlier study (1968) in light of the current study, it appears that the rate differences reported were unrelated to motivation. In addition to the differences in learning found for material presented at different rates, there is a possibility that the length of the segments heard, the presence and/or length of the pauses involved, or the total study time involved may have had a bearing on learning. Recent research has identified these factors as possibly influencing aural learning. Further research will be necessary to follow this up.

CHAPTER XXII

COMPREHENSION OF NARRATIVE PASSAGES BY FOURTH-GRADE CHILDREN AS A FUNCTION OF LISTENING RATE AND ELEVEN PREDICTOR VARIABLES

Robert L. Gropper*

Introduction

Since spoken language is the major mode of communication, an individual's ability to listen and to comprehend is of crucial importance. This skill, however, is not distributed equally among the population.

During the past 10 years great improvement in the technology associated with the recording and reproduction of speech has been made. One of the variables dealt with is the rate of presentation of speech.

Garvey (1953b) made the first attempt to alter the word rate of recorded speech. Although he employed a rather primitive manual method, his effort laid the groundwork for future progress. Fairbanks, Everitt, and Jaeger (1954) improved upon Garvey's work by electromagnetically discarding brief segments of recorded messages. Since that time, improved devices have been made available commercially. Two machines, the Tempo Regulator and Eltro Information Rate Changer, are presently in use. These devices permit the compression or expansion of speech without significant alteration of pitch or intelligibility.

Although progress has been made concerning the equipment and presentation variables of compressed speech, little has been done to examine the interaction of subject characteristics with performance at different levels of speech compression. The two factors which determine the efficiency of compressed speech are intelligibility and comprehension. The relationship of intelligibility to compressed speech is basic to the understanding of factors affecting comprehension.

---

*Dr. Robert L. Gropper is Assistant Professor of Education at the University of Miami, Coral Gables, Florida 33124.

Much research in the past has dealt with intelligibility. The usual index used to measure the intelligibility of time-compressed speech has been the ability of a subject to repeat brief messages with accuracy. The intelligibility score is the percentage of total words identified correctly. A summary of past research shows that, in general, increasing the amount of time compression appears to have a smaller influence on intelligibility than on comprehension.

Factors that affect comprehension include difficulty of the passage, vocal quality and style of the reader, and several unidentified listener variables. The usual procedure used for measuring comprehension is presentation of a tape-recorded message followed by a multiple-choice test over the contents. The results have been contradictory as to the influence of sex, IQ, chronological age, and passage difficulty. Research in the past has been done mainly with small samples using between-subjects designs. Past research has not lent itself to making predictions about the performance of individual subjects.

Purpose

The purpose of this study was to investigate the comprehension of narrative passages by fourth-grade children as a function of listening rate and performance on 11 predictor variables. These measures were evaluated with respect to their effectiveness in predicting levels of performance and learning efficiency indexes (learning per unit of time) across rates of speed. The narrative passages were presented at rates of 126, 190, 252, 312, and 380 words per minute (wpm). The criterion data were obtained from the Ss responses to multiple-choice questions based on the content of the passages. The intercorrelations were used to develop multiple regression equations. These equations were used to predict two aspects of performance at each rate: absolute level of performance (Performance) and the learning efficiency (Efficiency).

The results of this study may provide the basis for a useful technique of predetermining for individual pupils the most effective rate of orally-presented materials. The most efficient combination of predictors can be used to establish a standard for predicting performance across listening rates for individual Ss.

Method

This section contains a description of the Ss, instruments, apparatus and materials, and procedures used in the study.

Subjects

Three fourth-grade classes were used in this study. These classes were taken from schools in neighborhoods characterized by mixed socioeconomic levels. All children attending the classes were selected to participate. The total *S* pool consisted of 72 children. According to school records and teacher reports, all *Ss* possessed adequate auditory and visual acuity. The *Ss* in each classroom were randomly assigned to one of five experimental groups until a total of eight *Ss* per group was obtained. The remaining 32 *Ss* served as controls. None of the *Ss* had previous experience with compressed speech.

Instruments Used as Predictors

Each *S* was given a battery of nine tests designated as predictors. Tests were selected which appeared to measure components essential to performance with compressed speech materials. In addition to the tests administered by the *E*, a group IQ score, obtained from the *Ss*' cumulative records, and chronological age were used as predictors. The tests were administered to each *S* in the same order as they are listed below.

Intelligence. The Otis Quick-Scoring Mental Ability Test (Otis, 1954) is a group intelligence test. Scores for 37 of the 40 *Ss* were available in their cumulative folders.

Reading comprehension. The reading subtest of the Metropolitan Achievement Test (MAT) was used to test reading comprehension (Durost, Bixler, Hildreth, Lund, & Wrightstone, 1959).

Listening comprehension. Form B of the MAT Reading Subtest (Oral MAT) was presented to the *Ss* in a group at each of the three schools Durost et al., 1959). This test sampled the areas mentioned above; however, listening rather than reading comprehension was emphasized.

The passages and corresponding questions were recorded by the *E*. The *Ss* received an answer sheet containing 44 multiple-choice answer blanks. They were required to listen to each story and the questions and then to fill in the appropriate blanks. There was no visual contact with the passages. The passages and questions were recorded in about 22 minutes.

Perceptual speed. Perceptual speed is one of the five subtests from the Primary Mental Abilities Test, Grades 2-4 (Thurstone & Thurstone, 1963). This measures the ability to recognize quickly and accurately the likenesses and differences between objects or symbols.

Digit span. Digit span is a subtest taken from the Wechsler Intelligence Scale for Children (WISC) devised by Wechsler (1949). It is an auditory measure of short-term memory of digits presented sequentially.

Coding. Coding is a subtest of the WISC (Wechsler, 1949). The S must associate paired symbols and write the correct response element for a random series of stimulus elements.

Oral vocabulary. The Peabody Picture Vocabulary Test (PPVT) is an individually-administered test yielding an oral vocabulary score (Dunn, 1959).

Auditory attention span. The Auditory Attention Span for Related Syllables is a subtest of the Detroit Tests of Learning Aptitude (Baker & Leland, 1967). The test is a measure of short-term memory for sentences.

Clerical speed. Clerical speed and Accuracy is a subtest of the Differential Aptitude Tests (Bennett, Seashore, & Wesman, 1961). This test measures how quickly and accurately the S can compare letter and number combinations.

Auditory-vocal sequencing. The Auditory-Vocal Sequencing Test is a subtest of the Illinois Test of Psycholinguistic Abilities (ITPA) devised by McCarthy and Kirk (1961). The purpose of this test is to assess the S's ability to reproduce a sequence of auditory stimuli from memory. Approximately 5 minutes per S were necessary for completion.

Criterion tests. Each of the five criterion tests consisted of 22 to 27 multiple-choice items which sampled the retention of information from within the compressed speech material.

Apparatus and Materials

The apparatus used for this study included the following: (a) 5 sets of Calrad foam rubber, padded earphones, model HP4, of high impedence (15,000 ohms), stereo quality--1 set was used by the E and 4 sets were used by the experimental Ss; (b) 1 Wollensak magnetic tape recorder, model T-1500; (c) 1 junction box; (d) 1 "beep" box; (e) 1 slide projector, Kodak Carousel model 650; (f) 1 projection screen.

The Wollensak recorder was used with these settings: tone control on "HI-FI," volume control on "7," and tape speed on 7 1/2 ips. The junction boxes contained independently adjustable 50,000 ohm signal controls for each ear. A removable connecting cord was plugged into the tape recorder and into the junction box. Four sets of earphones were operated from the Wollensak by means of the junction box. The "beep" box was connected between the tape recorder and slide projector. It advanced the slides automatically by means of a recorded signal.

Materials used in this study included five Negro Heritage passages and associated tests. These passages were originally prepared for use in a study by Woodcock and Clark (1968c). The contents of these passages concerned the lives of Estevenico, Jesse Owens, J. B. Olinger, Harriet Tubman, and William C. Handy.

The times of the five passages at the original recorded rate of about 190 wpm were: "Harriet Tubman," 21 minutes; "J. B. Olinger," 14 minutes 20 seconds; "Estevanico," 19 minutes 10 seconds; "Jesse Owens," 19 minutes 25 seconds; and "W. C. Handy," 16 minutes 20 seconds. Upon rerecording at compressed or expanded rates, the listening times were changed proportionately.

Each tape consisted of the following: (a) instructions to the S regarding the earphones and adjustment of volume to each ear; (b) instructions regarding the listening task to be presented; (c) the passage, at the appropriate wpm rate; (d) instructions for taking the test; (e) the multiple-choice test covering the contents of the passage.

The tapes used with the control groups, who did not listen to the passage did not include sections (b) and (c) above. All instructions and tests wer presented at the normal speech rate on the tapes; the passage section wa the only portion which was compressed or expanded.

Procedure

The sequence of steps in developing and standardizing criterion tests, selecting Ss, administering the predictor tests, and conducting the experimental sessions is outlined below.

Developing criterion tests. The 20 existing questions for each of 20 Negro Heritage passages were examined. The five passages containing the items which best discriminated the Ss who made scores in the upper half of the class from those in the bottom half were selected. These passages were read carefully and about 30 new multiple-choice questions written. Pupils from three fourth-grade classes listened to the passages at normal speeds and took the multiple-choice tests. These items were analyzed by means of a tetrachoric coefficient to find the item discrimination indexes. The passages and tests were presented by tape recorder to an entire class. Approximately 30 minutes were needed to listen to and be tested over one passage. Five sessions were needed with each class to complete the procedure for all five stories.

Standardizing criterion tests. The 30 new items for each passage were combined with the 20 previous items. These pooled questions were examined and those with a discrimination index of .5 or better were selected. Each of four new classes as a group listened to the five passages and associated slides at normal speeds and took the tests. The passages and the questions were presented by means of tape recorder. After all tests were scored, a normalized T score transformation was applied and the scores were used for the experiment.

Selecting subjects. Three fourth-grade classes from a mixed socioeconomic school district were selected to participate. The Ss were randomly assigned to groups. All students not selected from the three classes served as controls.

Administering the predictor tests. All predictors with the exception of the Otis Intelligence Test were administered by the E. The MAT Reading subtest and the Oral MAT were administered to the entire class at each school on the first day of testing. During the second session, Perceptual Speed and Clerical Speed were given to the experimental Ss in groups of four. All testing from this point was done where the experimental apparatus was arranged.

During the third session each S individually took the Digit Span and Coding subtests of the WISC, followed by the PPVT. At the last sessions, the Ss were given the Auditory Attention Span subtest and the Auditory Vocal Sequencing test. The tests were arranged in this manner so that no S would spend more than 20 minutes in the testing situation. A fifth and sixth session were necessary to make up for Ss who were absent during the second, third, or fourth sessions.

Conducting the experimental sessions. The day after all Ss had been given the predictor tests, experimental groups were brought to the room of their school where the experimental apparatus had been arranged. Each group heard one story per session. Figure 22.1 shows the order of stories and speeds for each group. The Ss were told that they would listen to a story through the earphones and watch slides on a felt board about 4 feet in front of them. They were shown their volume controls and instructed to put on the earphones.

A familiarization tape was played for 3 minutes at the same speed as the experimental tape. Immediately following the removal of the training tape, the experimental tape was used. From this time to the end of each session, all instructions were contained on the tape.

Groups	126	190	252	312	380
I	1st Estevanico	2nd Handy	3rd Owens	4th Tubman	5th Olinger
II	2nd Handy	3rd Estevanico	5th Olinger	1st Owens	4th Tubman
III	3rd Owens	5th Tubman	4th Estevanico	2nd Olinger	1st Handy
IV	4th Tubman	1st Olinger	2nd Handy	5th Estevanico	3rd Owens
V	5th Olinger	4th Owens	1st Tubman	3rd Handy	2nd Estevanico

Rate in wpm

Figure 22.1. Presentation sequence by rate and passage for each group.

The listening time spent with each passage ranged from 7.27 to 31.02 minutes depending upon the rate of compression. Table 22.1 shows the amount of time spent with the passages at each speed.

TABLE 22.1

THE TIME IN MINUTES BY SPEED FOR EACH PASSAGE

	126 WPM	190 WPM	252 WPM	312 WPM	380 WPM
Estevanico	28.80	19.20	14.40	11.52	9.60
Olinger	21.80	14.53	10.90	8.72	7.27
Tubman	31.02	20.68	15.51	12.41	10.34
Owens	29.00	19.33	14.50	11.60	9.67
Handy	24.35	16.25	12.19	9.75	8.13

Following the passage and during the presentation of instructions for the tests, the E distributed a copy of the questions and a pencil to each S. The test items were presented at normal speeds simultaneously on the tape as the Ss followed on their printed test form and selected answers. At the completion of the test, the Ss were instructed to remove their earphones. The entire listening time, from putting on earphones to taking off earphones, was approximately 8 minutes more than the listening time for the passage. Thus, the total time varied from 15.3 minutes to 39 minutes.

The control groups followed the same procedure except that their taped instructions went directly from adjustment of earphones into the test instructions and the test. Any reference to the listening passage and the passage itself was bypassed.

The same procedure was followed for five sessions. The criterion tests were then scored by the E. Raw scores were converted into normalized T scores using the norms provided in the pilot study.

Statistical analyses. The R01: Regression Analysis for Raw Data Program at the Peabody Computer Center was used to compare the predictor and criterion tests. The program computes means, standard deviations, Pearson product-moment correlations, and regression equations. The printed output consists of means, standard deviations, and a correlation matrix. B weights, a regression constant, and an iteration sequence are also provided for each regression equation.

Appropriate F ratios (Walker & Lev, 1953) were used to test the relations between speeds for absolute level of performance and learning efficiency. A Newman-Keuls procedure (Winer, 1962) was used for multiple t comparisons following significant main effects. The .01 level of significance was employed to evaluate the statistical significance of all comparisons.

Results

The results of the analyses of the test data are reported herein. All results presented in connection with this investigation are reported on the basis of the scores of the 40 Ss receiving the 10 predictor and 10 criterion tests.

Descriptive Data

Table 22.2 contains means and standard deviations for scores on each of the variables under consideration. The PPVT score is reported in months representing "receptive language age." The MAT Reading test is reported as a standard score using the tables provided. The remaining predictor tests with the exception, of course, of IQ are presented as raw scores.

The criterion scores are reported as normalized T scores. The Learning Efficiency scores were computed using the following formula:

$$\text{Learning Efficiency Index} = \frac{\text{Treatment Mean--''Test Only'' Mean}}{\text{Listening Time in Minutes}}$$

"Test Only" refers to the control group who took the multiple-choice criterion tests in the same manner as the experimental group, but did not have the benefit of listening to the stories.

Correlation Analyses

The main concern of this investigation was to obtain multiple regression equations based on the correlations of the predictor tests with each of the criterion scores.

A sequence of iterations were computed in which the predictor variables were listed in the order of the amount of variance they explained. Tables 22.3 and 22.4 list the iteration sequence for each criterion for absolute level of performance and learning efficiency. The sequence was terminated when the additional predictors accounted for less than 2% of the variance.

TABLE 22.2

MEANS AND STANDARD DEVIATIONS FOR PREDICTORS AND CRITERION TESTS

Variable	Unit of Measure	$\overline{X}$	SD
Chronological Age	Months	114.80	3.74
PPVT (Receptive Language Age)	Months	115.80	17.51
Digit Span	Raw Score	8.95	1.58
Coding	Raw Score	39.50	7.88
Auditory Vocal Sequencing	Raw Score	28.30	6.42
Auditory Attention Span	Raw Score	50.47	12.62
Perceptual Speed	Raw Score	21.82	11.74
Clerical Speed	Raw Score	66.95	18.28
MAT Reading	Standard Score	56.07	6.65
Listening Comprehension	Raw Score	17.55	7.67
Otis IQ	IQ	104.47	11.08
Criterion Tests: Performance			
126 WPM	T Score	54.33	6.81
190 WPM	T Score	52.73	7.86
252 WPM	T Score	51.24	7.61
312 WPM	T Score	44.27	9.18
380 WPM	T Score	39.49	8.62
Criterion Tests: Efficiency			
126 WPM	Efficiency Index	.85	.31
190 WPM	Efficiency Index	1.19	.51
252 WPM	Efficiency Index	1.43	.67
312 WPM	Efficiency Index	1.12	.90
380 WPM	Efficiency Index	.87	.98

TABLE 22.3

THE ITERATION SEQUENCE FOR ABSOLUTE LEVEL OF PERFORMANCE

Criterion	Predictors	Cumulative r^2
126 WPM	MAT	.22
	Digit Span	.29
	CA	.35
	Perceptual Speed	.40
	Auditory Attention Span	.42
190 WPM	MAT	.37
	Digit Span	.48
	Auditory Vocal Sequencing	.55
252 WPM	PPVT	.18
	Auditory Vocal Sequencing	.20
	CA	.22
	Clerical Speed	.26
326 WPM	MAT	.28
	Digit Span	.37
	CA	.44
380 WPM	Listening Comprehension	.24
	PPVT	.36
	Clerical Speed	.38
	Digit Span	.40

Multiple Regression Analyses

Using the predictors listed in Tables 22.3 and 22.4, the data was re-analyzed in order to obtain the B weights and the constant for multiple regression equations. These equations, listed in Tables 22.5 and 22.6, give the formulas for using the best set of predictors at each speed for absolute level of performance and learning efficiency.

The major purpose of this study was to identify a reliable set of predictors that could be used with all speeds. Four predictors were selected for this purpose. A point system was devised to select the tests to be used in the final battery. Table 22.7 graphically indicates the iteration sequence for all speeds and points assigned to each test.

TABLE 22.4

THE ITERATION SEQUENCE FOR LEARNING EFFICIENCY

Criterion	Predictors	Cumulative r^2
126 WPM	PPVT	.18
	Perceptual Speed	.23
	MAT	.28
	PPVT	.32
	CA	.37
	Auditory Attention Span	.41
190 WPM	MAT	.22
	Digit Span	.30
	Auditory Vocal Sequencing	.38
252 WPM	PPVT	.15
	IQ	.18
	Clerical Speed	.23
	Auditory Vocal Sequencing	.30
	Coding	.34
312 WPM	MAT	.29
	PPVT	.35
	CA	.38
	Digit Span	.41
380 WPM	Listening Comprehension	.33
	PPVT	.40

On the basis of the results in Table 22.7, the MAT Reading, the PPVT, Digit Span, and Listening Comprehension test were selected as the battery of tests for use as a predictor of performance at all speeds. This battery takes approximately 55 minutes to administer. Table 22.8 shows the multiple regression equations at each speed for absolute level of performance. The amount of variance accounted for is also indicated. Table 22.9 gives the same information for the learning efficiency speeds.

Analysis of the Criterion Scores

Figures 22.2 and 22.3 show the curves generated by plotting the _T_ scores for Performance and Efficiency. A comparison of the two graphs shows that, though Performance becomes progressively poorer with an increase in speed, Efficiency peaks at the middle speed.

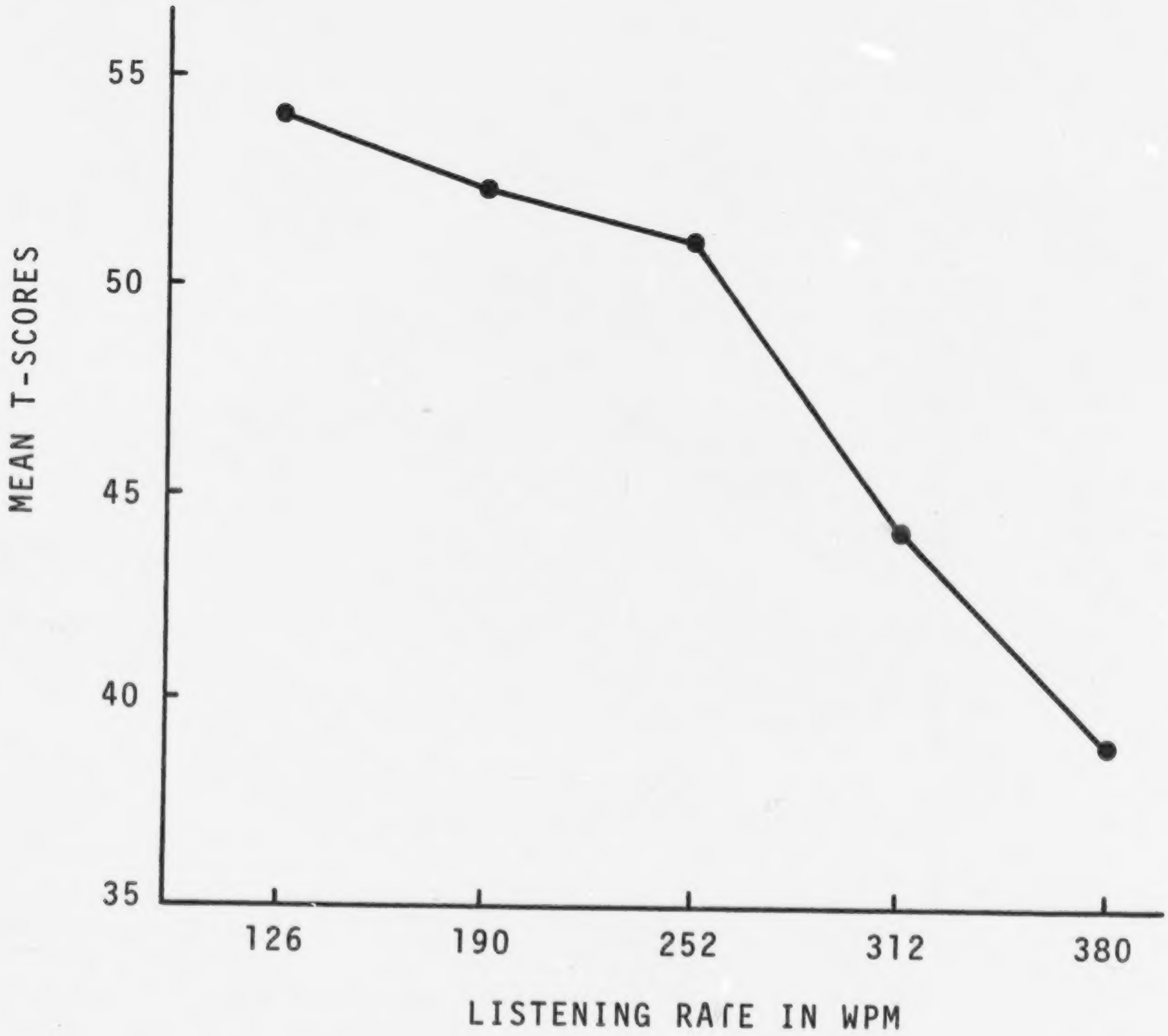


Figure 22. 2. T score means for performance.

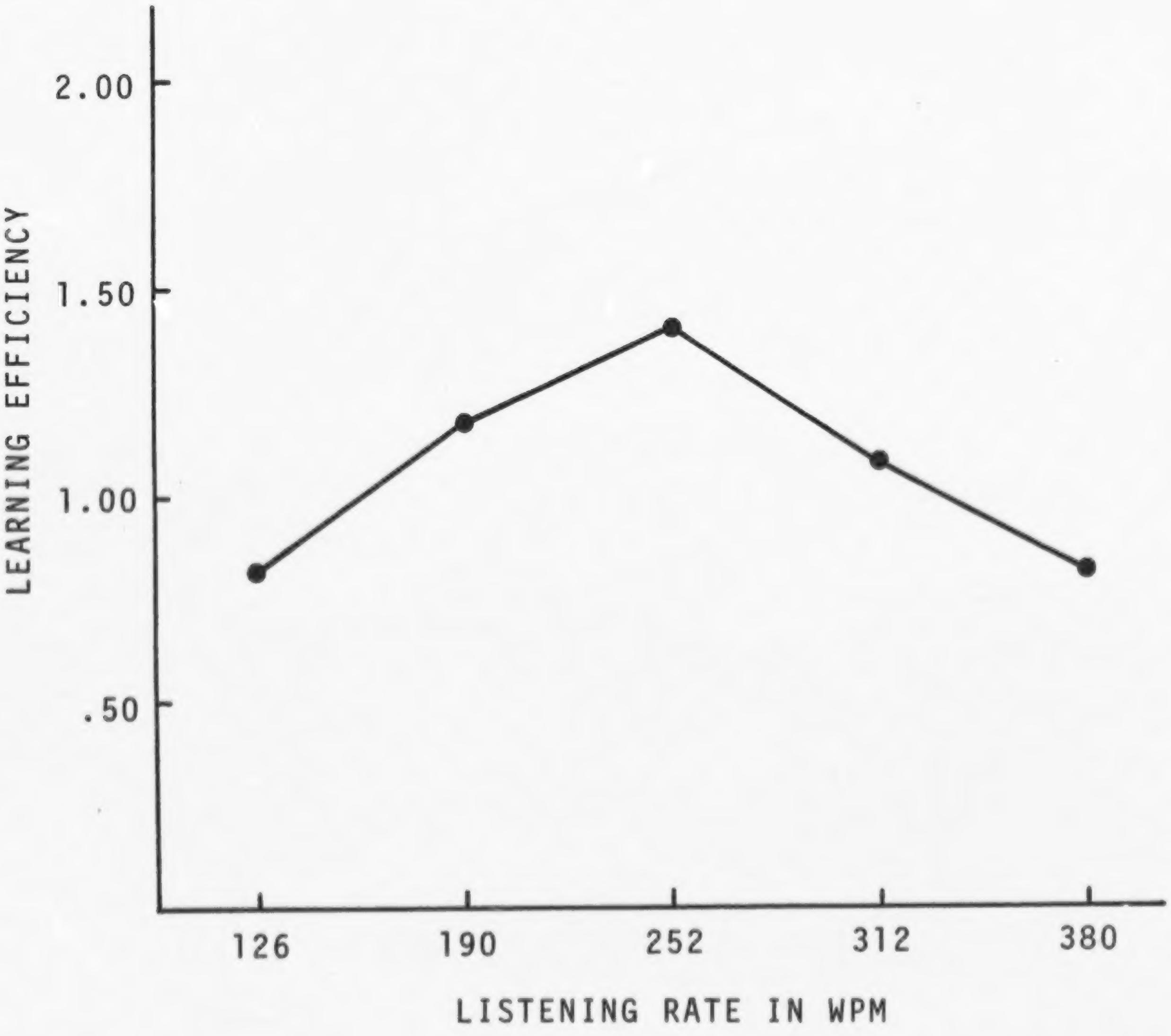


Figure 22.3. Mean efficiency index for each rate.

TABLE 22.5

MULTIPLE REGRESSION EQUATIONS FOR ABSOLUTE LEVEL OF PERFORMANCE

Criterion	Constant	B Weights (Predictor)	R	R^2
126 WPM	-44.31	+.66(MAT) +1.62(Digit Span) +.50(CA) -.19(Perceptual Speed) -.13(Auditory Attention Span)	.65	.42
190 WPM	.91	+.72(MAT) +2.70(Digit Span) -.36(Auditory Vocal Sequencing)	.74	.55
252 WPM	25.7	+.13(PPVT) -.25(Auditory Vocal Sequencing) +.22(IQ) -.09(Clerical Speed)	.52	.27
312 WPM	-84.8	+.59(MAT) +2.18(Digit Span) +.66(CA)	.66	.44
380 WPM	20.42	+.52(Listening Comprehension) +.18(PPVT) -.08(Clerical Speed) -.77(Digit Span)	.63	.40

TABLE 22.6

MULTIPLE REGRESSION EQUATIONS FOR LEARNING EFFICIENCY

Criterion	Constant	B Weights (Predictor)	R	R^2
126 WPM	- 2.87	+.01(PPVT) -.01(Perceptual Speed) +.03(MAT) +.01(CA) +.01(Auditory Attention Span)	.65	.42
190 WPM	- 1.34	+.04(MAT) +.17(Digit Span) -.03(Auditory Vocal Sequencing) +.01(Coding)	.64	.41
252 WPM	- 1.43	+.005(PPVT) +.02(IQ) -.01(Clerical Speed) -.04(Auditory Vocal Sequencing) +.15(Digit Span) +.01(Auditory Attention Span)	.59	.35
312 WPM	-10.23	+.05(MAT) +.01(PPVT) +.05(CA) +.12(Coding)	.65	.42
380 WPM	- 2.06	+.06(Listening Comprehension) +.02(PPVT)	.63	.40

TABLE 22.7

THE POINT VALUES ASSIGNED TO THE PREDICTORS ON THE BASIS OF POSITION IN THE ITERATION SEQUENCES

Variables	1st (4 Pts.)	2nd (3 Pts.)	3rd (2 Pts.)	4th or Lower (1 Pt.)	Total
MAT Reading	5		1		22
PPVT	3	3			21
Digit Span		4		2	14
Listening Comprehension	2		1		10
Auditory Vocal Sequencing		1	2	1	8
IQ		1	1	1	6
Clerical Speed			2	1	5
Perceptual Speed		1		1	4
Auditory Attention Span				3	3
CA			3		3
Coding				2	2

TABLE 22.8

MULTIPLE REGRESSION EQUATIONS FOR PERFORMANCE USING FOUR SELECTED PREDICTORS

Criterion	Constant	B Weight (Predictor)	R	R^2
126 WPM	17.80	+.06(MAT) +.99(PPVT) +.40(Digit Span) -.15(Listening Comprehension)	.57	.33
190 WPM	5.09	+.07(MAT) +1.53(PPVT) +.41(Digit Span) +.15(Listening Comprehension)	.71	.51
252 WPM	24.98	+.14(MAT) +.40(PPVT) +.09(Digit Span) +.03(Listening Comprehension)	.44	.20
312 WPM	5.00	+.10(MAT) +1.60(PPVT) +.36(Digit Span) +.16(Listening Comprehension)	.64	.41
380 WPM	17.42	+.21(MAT) -.81(PPVT) -.06(Digit Span) +.47(Listening Comprehension)	.61	.38

TABLE 22.9

MULTIPLE REGRESSION EQUATIONS FOR EFFICIENCY USING FOUR SELECTED PREDICTORS

Criterion	Constant	B Weights (Predictor)	R	R^2
126 WPM	- .52	+.004(MAT) +.004(PPVT) +.017(Digit Span) -.011(Listening Comprehension)	.52	.27
190 WPM	- 1.05	+.004(MAT) +.08(PPVT) +.01(Digit Span) +.02(Listening Comprehension)	.61	.37
252 WPM	- .55	+.01(MAT) +.07(PPVT) -.005(Digit Span) +.004(Listening Comprehension)	.43	.18
312 WPM	- 3.82	+.01(MAT) +.08(PPVT) +.05(Digit Span) -.001(Listening Comprehension)	.61	.37
380 WPM	- 1.55	+.02(MAT) -.06(PPVT) -.007(Digit Span) +.07(Listening Comprehension)	.64	.41

A one-way analysis of variance was performed to compare speeds. Table 22.10 presents the results of the analysis for absolute level of performance. The difference between speeds was significant at the .01 level of confidence.

TABLE 22.10

ANALYSIS OF VARIANCE: PERFORMANCE

Source	df	MS	F	p
Speeds	4	1589.968	24.062	.0001
Error (G)	195	66.076		
Total	199			

The Newman-Keuls test was used to determine the significance of difference among means. Table 22.11 summarizes the results. All comparisons which included the two fastest speeds were significant.

TABLE 22.11

NEWMAN-KEULS TEST: PERFORMANCE

Ordered Means	312 WPM A_4	252 WPM A_3	190 WPM A_2	126 WPM A_1
380 WPM A_5	4.78**	11.75**	13.12**	14.97**
312 WPM A_4		6.97**	8.44**	10.19**
252 WPM A_3			1.37	3.22
190 WPM A_2				1.85
Critical Values ($p = .01$)	$r = 2$ 6.02	$r = 3$ 6.79	$r = 4$ 7.26	$r = 5$ 7.59

Table 22.12 presents the summary statistics for the one-way analysis of variance performed to compare speeds for Efficiency. The differences between speeds was significant to the .01 level of confidence.

TABLE 22.12

ANALYSIS OF VARIANCE: EFFICIENCY

Source	df	MS	F	p
Groups	4	2.44	4.59	.001
Error	195	.53		
Total	199			

The Newman-Keuls test was also performed to determine the significance of differences among means for Efficiency scores. The results, which appear in Table 22.13, show that the 126 wpm versus 252 wpm and 380 wpm versus 252 wpm comparisons were significant to the .01 level of confidence.

Discussion

The data derived from this study were analyzed in the following ways: (a) the predictor tests were correlated with the criterion tests over the contents of the passages; (b) multiple regression equations using the coefficients were obtained; (c) analysis of variance was performed on the criterion scores; (d) individual means were compared by the Newman-Keuls procedure. Each of the above analyses yielded information which may influence future studies of aural comprehension. The discussion which follows describes the population from which the data were derived, discusses each of the analyses described above, and introduces implications for future consideration.

TABLE 22.13

NEWMAN-KEULS TEST: EFFICIENCY

Ordered Means	380 WPM A_4	312 WPM A_3	190 WPM A_2	252 WPM A_1
126 WPM A_5	.00	.27	.35	.58**
380 WPM A_4		.25	.33	.56**
312 WPM A_3			.08	.31
190 WPM A_2				.23
Critical Values ($p = .01$)	$r = 2$ .40	$r = 3$ .45	$r = 4$ .48	$r = 5$ .51

The Sample

The experimental Ss in this study were 40 fourth-grade children from several socioeconomic backgrounds. Examination of these Ss' performance yielded descriptive information. Scores on the PPVT, MAT Reading, Coding, Digit Span, Otis IQ, and Perceptual Speed tests indicated that as a group the Ss were performing at about an average level in terms of grade expectancy. Four additional measures were included. These were Listening Comprehension, Clerical Speed, Auditory Vocal Sequencing, and Auditory Attention Span. No age norms or grade equivalents were available for these measures. Raw scores from these tests were used in developing the predictive indexes described later. On the basis of performance on the predictor tests mentioned above, the Ss can be described as average with individual variation ranging from educable mentally retarded to superior levels.

Correlational Analyses

Correlational analyses were performed in an effort to identify a relatively short battery of tests which could predict comprehension for an individual S exposed to compressed and expanded speech. The analysis utilized the multiple correlation coefficient derived from scores on 11 predictor measures and criterion tests. The criterion tests were 10 multiple-choice tests administered to Ss after they had listened to passages presented at different rates of speed.

The simple correlations of predictor tests with criterion scores ranged from -.17 to +.61. These correlations included both Performance and Efficiency criterion tests. Most of the predictors were correlated fairly evenly across speeds. A notable exception was the coefficient obtained for the Listening Comprehension test. Listening Comprehension was a poor predictor at three speeds and a moderately poor predictor at a fourth speed. However, at the fastest speed, 380 wpm, Listening Comprehension was clearly the best predictor. This may be explained by the fact that both performance on the Listening Comprehension test and performance on the criterion tests for 380 wpm require a great deal of aural concentration. The Listening Comprehension test served as a comparatively strong predictor of success with compressed speech at speeds near 380 wpm, being associated with 25 and 35% of the variance in Performance and Efficiency respectively.

MAT Reading and the PPVT were correlated with criterion tests at .45 and .44 respectively, using data for all listening rates. These results can be compared with those of Condon (1965), Duker (1965), and Fawcett (1965) who found mean correlations ranging from .47 to .74 between performance on reading and listening at normal rates.

A correlation of .60 was obtained between MAT Reading and the PPVT. This amount of overlap reduced the individual contributions of these tests when included in a predictor battery.

Digit Span correlated moderately with the criterion tests at every speed, with the exception of the fastest presentation rates (380 wpm). The mean correlation of Digit Span with scores obtained at all slower rates was .36. The fact that Digit Span correlated relatively poorly with MAT Reading (.27) and the PPVT (.38) made it an efficient addition to the predictor battery. These results are in agreement with the fact that historically Digit Span has had a poor correlation with overall performance of the WISC. While not being highly correlated with intellective factors, Digit Span seems to be a good test of short-term memory and listening comprehension.

The remaining measures were associated with very little of the variance at each speed when added to the predictor battery. Therefore, the group of tests discussed above was selected as the final predictor battery. The variance associated with the Performance criterion accounted for by these tests ranged from a low of 20% at 252 wpm to a high of 51% at 190 wpm. The variance associated with the Efficiency criterion ranged from 18% at 252 wpm to 41% at 380 wpm.

At the 252 wpm rates, the amount of variance associated with predictors can be improved by substituting Auditory Vocal Sequencing, Coding, and Clerical Speed for MAT Reading, Listening Comprehension, and Digit Span. This substitution can raise the common variance by 7 and 18% for Performance and Efficiency respectively. At the remaining speeds, the original battery is generally the most effective.

An examination of the intercorrelations of the five criterion tests for Performance indicates that in some cases results at one speed may be a good predictor of performance at a second speed. Results for the 126 wpm rate correlated above .50 with all but the 380 wpm rate. In addition, 190 wpm and 312 wpm correlated .50 or above as did 252 wpm and 380 wpm.

The mean correlation among Performance criterion tests was .497. The 126 wpm rate had a .54 mean correlation with the remaining speeds. This was the highest mean correlation. These results may seem unusual in that one would expect the rates closest together to have the highest correlations with each other. However, this is not the case, since the 126 wpm rate and the 252 wpm rate were highest correlated. These rates were two steps removed from each other on the continuum defined in this study. Also, the 312 wpm rate correlated highest with the 252 wpm rate, which similarly is two steps away. The only adjacent speeds which show high correlations relative to other speeds are 190 wpm and 252 wpm. More within-Ss data are needed before definitive statements can be made as to the predictive strength of performance at one speed when predicting performance at another speed.

The intercorrelations of the five speeds for Efficiency were much lower. This was expected because of the formula used to obtain Efficiency scores. The mean intercorrelation was .35. An unexpectedly high correlation of .68 occurred between the 126 and 312 wpm rates. The difference was .24 between this and the second highest correlation. If this coefficient is reliable, then these rates are better predictors of each other than is the battery of tests identified above. Past research has not investigated intercorrelation between speeds using a within-Ss design. This design needs to be replicated and the results compared to see if intercorrelations among speeds are high enough to permit reliable predictions to be made across speeds.

Analysis of Variance

A secondary purpose of this study was to examine the criterion scores to find the best presentation rates in terms of Performance and Efficiency. As speed increased, Performance on the criterion tests decreased. An analysis of variance showed that the difference among presentation rates was significant. A Newman-Keuls procedure indicated that significance was obtained only when comparing speeds above 252 wpm. In other words, a significant drop in performance did not occur until the S listened to speeds faster than 252 wpm. This is in basic agreement with past research (Bixler, Foulke, Amster, & Nolan, 1961; Foulke, 1966a; Woodcock & Clark, 1968a, 1968b).

An examination of the Efficiency curve (see Figure 22. 3) revealed that the 252 wpm rate fared best in terms of learning per unit of time. This rate is about 25% faster than normal. This was in agreement with the findings of Woodcock and Clark (1968b). An analysis of variance showed that the difference between speeds was significant. Individual comparisons, however, revealed that only the 252 wpm rate versus the 126 wpm rate and 380 wpm rate comparisons were significant. Large individual differences were obtained, indicating that there is not one most efficient speed for everyone. In most cases, however, a speed much slower than normal will not add much to comprehension, while speeds about twice as fast as normal will take too much away from comprehension to warrant their use.

Implications for Future Research

This study represented an initial attempt to use a within-Ss design to identify an efficient set of predictors of performance with compressed speech materials. Cross validation procedures are necessary. Populations of high-, average-, and low-IQ Ss should be used to get supplementary data on the present battery of predictors. Perhaps new tests can be added or substituted to refine the initial predictor group. A more reliable predictor is needed for speeds around 252 wpm, especially since this is the most efficient speed for Listening Comprehension.

More materials have to be adapted and standardized so that future research may use a within-Ss design. In this way, more speeds may be used while not limiting the number of Ss per cell per speed.

Finally, past research has yielded little knowledge regarding the nature of listening tasks and of training methods for promoting listening skills. Studies attempting to correlate existing measures with comprehension of accelerated speech may have the overriding impact of identifying factors both cognitively and perceptually involved in the operation of listening.

CHAPTER XXIII

A COMPARISON OF TWO TECHNIQUES FOR INCREASING THE RATE OF READING OF SIXTH-GRADE GIFTED PUPILS: THE COMPRESSED SPEECH MACHINE AND THE SELF-IMPROVEMENT METHODS

Charles R. Walker*

Design of the Study

Selection of Subjects

The 71 accelerated sixth-grade students in the Centennial Schools were chosen as the Ss of the experiment. These students will most likely continue their formal education after high school and will probably have the greatest need to increase their rate of reading because of the demands placed upon them. Children in Centennial are identified for the gifted program either in the beginning of fourth grade or sixth grade. All of these children have been administered an individual intelligence test, usually a Stanford-Binet or a Wechsler Intelligence Scale for Children (WISC), and must have obtained a score of 130 or more to qualify for the program. In addition to ability, consideration is given to such factors as achievement, maturity, teacher recommendation, etc. Prior to the beginning of school, the Director of Elementary Education grouped these students into four classes on the basis of past achievement, intelligence, and other available objective data.

Establishment of Experimental Groups

At the beginning of the experiment the entire 71 pupils were administered a rate pretest. The pretest consisted of having the pupils read silently for 15 minutes from a Landmark book randomly selected from a total of 105 books. The Landmark Books are a series of high interest nonfiction stories adapted for young readers. The books are basically sixth-grade readability level and these were the books which were read in the project by both experimental groups. The books read in the project again were randomly selected. At the end of 15 minutes, reading was stopped and the pupils drew a line after the last word they had read. The examiner established a distribution of scores and determined the words per minute (wpm) rate for each child. On the basis of sex and pretest rate--fast and slow as determined by the common median of each sex pretest rate score distribution--stratified randomization was used. Taking one stratum at a time, the pupils were randomly assigned to one of the three groups (compressed speech group, self-improvement group, or control group). Because of randomization, one can assume that all three groups were initially equal in all important respects.

---

*Charles R. Walker is Assistant Superintendent for Elementary Education, Centennial Schools, Warminster, Pennsylvania 18974.

Description of the Tests Used

Prior to the beginning of the study itself and again at the conclusion of the study, the three groups were given Gates Reading Survey--grades 3 to 10 (1960 reprint of 1958 edition). This is a nationally recognized, standardized test that provides evaluation in three fundamental processes of reading: a 65-item Vocabulary subtest, a 21-item Comprehension subtest, and a 36-item Speed and Accuracy subtest. The Vocabulary subtest is not a speed test. Pupils were allowed as much time as they needed to complete the questions. The Comprehension subtest consisted of 21 passages arranged in order of difficulty. The influence of speed in reading was eliminated by allowing the pupils as much time as they required. The Speed subtest consisted of 36 paragraphs of basically equal difficulty. Each one contained a comprehension exercise to determine whether it was understood. The time given was limited, with the result that the score--number of exercises correct--represented the speed in reading. The Gates test was used because it provides a range of ability and norms extending from 1.5 to grade 13 which was necessary with sixth-grade gifted pupils. Three equivalent forms exist: Form 1, Form 2, and Form 3.

Spache (1965) claims: "The reliability coefficients for each of the subtests . . . are in the .80's and are certainly adequate for most testing purposes [p. 1066]." To overcome order effects of the pretest-posttest administered, Forms 1 and 2 were randomly assigned to pupils for the pretest and then automatically each pupil received the alternate form as the posttest. This allowed for an unbiased analysis of gain scores.

The Landmark immediate rate posttest was given to all three groups at the conclusion of the 6 week experimental period. All pupils read silently for 15 minutes from one of the Landmark books randomly selected for that day. Again this randomization process ensured representative sampling of reading materials. At the end of 15 minutes, the pupils marked where they were and the examiner scored the tests for each group and established a rate (wpm) for each pupil.

A Landmark delayed rate posttest was given 6 weeks after the immediate rate posttest. The test procedure used was the same as the immediate rate posttest except for a different selection from the Landmark Books.

Daily Schedule for the Three Groups

Each morning all three groups received their regular reading instruction, which consisted of a basal program involving phonics, word attack skills, comprehension, vocabulary, etc. During the period of the experiment, which was 6 weeks, this regular reading program was continued under the direction of the 2 two-teacher teams. The regular reading instruction occurred during the morning. The experimental period, of 15 minutes each day, was also in the morning and scheduled the same time for all three groups. The experimental period was utilized as follows.

The Preston-Botel group. This group spent this period reading selections from the Landmark Books. These selections were mimeographed and sufficient in length so that no pupil could finish a selection within the 15 minute period. The technique which was used to attempt to increase the rate of reading was a commonly accepted one which is currently used in many schools. The technique has been described by Preston and Botel (1967, pp. 37-39). Basically, it is a structured self-improvement program in which the reader attempts to do the following: (1) select easy materials, (2) preview, (3) time himself and force himself to read more rapidly, and (4) keep a graphic record of his performance. Before the actual experiment began, the teachers explained this technique carefully to the pupils. Daily, at the conclusion of each experimental period, pupils in the Preston-Botel group were given an informal comprehension check. These checks consisted of a 10 question multiple-choice objective test covering the basic passage they had just completed. The purpose of the daily comprehension check was to make the pupils aware of the importance of reading with understanding. It is important to note that both the Preston-Botel group and the compressed speech group were reading the same Landmark Books selections and were given the same daily comprehension checks. This, besides the randomization procedures, ensured comparability of the two methods.

The compressed speech group. This group consisted of four subgroups as were determined by the results of the rate pretest (using the Landmark Books Series). These subgroups, during the 15 minute experimental period of each day, listened to tapes which had been prerecorded and compressed according to the reading rates established by the rate pretest (using the Landmark Books Series). As they listened to the tapes, they read silently the same passages from the Landmark Books.

The purpose for having the subgroups were twofold. The investigator wanted to tailor the program as much as possible to each individual. This is why he was listening to a tape recorded very close to his actual reading rate at that time. This enabled each pupil to begin at a point which was comfortable to him and to experience success. Having subgroups also aided in the management of the experiment. The investigator had available four tape recorders, each with six to eight sets of headphones. The compressed speech group had the four recorders assigned to it. The initial tape, listened to by each of the subgroups, was played at a speed so that the members of each respective subgroup would have no difficulty in comprehending the material. This enabled all groups to acclimate to the recorders, use of headsets, and general procedures. There were 30 master tapes--each recorded at a uniform number of wpm and each containing different selections from the Landmark Books with sufficient length so that the fastest reader would have a minimum of 15 minutes of material. These master tapes were compressed or decompressed to satisfy the requirements of each of the subgroups. The master tapes were made by an outside person who is an excellent reader and who had no connection or familiarity with the project. The Eltro Information Rate Changer was used to compress or decompress the tapes. All pupils in the compressed speech group were listening to and reading silently the same story but at varied rates.

When to compress the tape further and to what degree was determined by the daily in-process comprehension checks and the informal daily feedback from the pupils and teachers. The purpose of these checks was to make the pupils aware of the fact that they should always be reading for understanding and, that if understanding does not occur, then there is no value in increasing ones' rate. The check also served as a motivational tool.

Another purpose for their use was to serve as a check with the compressed speech group to determine the limits of compression--or, in other words, to what degree a pupil could have the passages speeded and still maintain a certain level of comprehension.

The daily check consisted of a 10 question multiple-choice objective test covering the basic passage they had just completed. The test questions were developed by an impartial outside person who had no connection with the project but who is a reading teacher and knows the kinds of questions relevant to comprehension.

However, it should be noted that these tests were simply an in-process check and were not used for statistical analysis but rather only as a procedural adjunct.

After the subgroups within the compressed speech group had listened to a series of tapes recorded at the same speed and had become acclimated to the technique, the instructor shut off the headsets and had the pupils continue reading without the auditory support. This motivational technique was done during the fifth week of the study. After 1 to 5 minutes, the pupils marked where they were, and the headsets were turned on. If, in fact, pupils were increasing their rate, they would be at about the same place in the selection as was the tape. This technique began to show evidence whether or not a pupil's rate really was increasing. The auditory support was turned off for periods of time ranging from 1 minute up to 5 minutes over a period of 5 days to see if the increased rate remained constant.

The control group. During the experimental period, the control group was doing independent reading. These would be normal reading activities involving the various disciplines or library books. The reading for this group was completely unstructured except for the 15 minute time limit which applied to all groups. Pupils were permitted to choose their own reading materials.

Teachers' Rotation Schedule

During the experiment, there was a rotating random schedule for the four teachers involved with the compressed speech group, the self-improvement group, and the control group. This prevented contamination which might have arisen as a result of differences in personalities and techniques among the teachers. This rotational scheme spread the effects of possible teacher difference in equal degrees throughout the three methods groups. For the first 4 days, for the four teachers, a 4 x 4 randomized rotation schedule was devised. This schedule was repeated systematically every 4 days throughout the experiment. Because there were only three groups, each teacher had a completely free period every fourth day.

Statistical Analysis

Because of the randomization method used, initial differences among groups can be considered negligible. Thus, regular analyses of variance were used. The three factors in each analysis of variance were methods, sex, and preexperimental reading rate. Eight separate analyses were performed: (a) Landmark immediate rate posttest, (b) Gates Survey

immediate posttest--Vocabulary, (c) Gates Survey immediate posttest--Speed, (d) Gates Survey immediate posttest--Comprehension, (e) Landmark delayed rate posttest, (f) reading rate gain scores (Gates Survey posttest minus Gates Survey pretest)--Vocabulary, (g) reading rate gain scores (Gates Survey posttest minus Gates Survey pretest)--Speed, and (h) reading rate gain scores (Gates Survey posttest minus Gates Survey pretest)--Comprehension. Also, these eight analyses of variance provided information on any difference that existed between sexes and between preexperimental reading rates. If in any of the analyses, overall significance of methods is found, the question of which particular method or methods caused the significant difference was answered by using Scheffe's technique of multiple comparisons.

Results

In reporting results the investigator chose the .05 level of significance. On the Landmark immediate posttest, significance was found in methods ($.01 < p < .025$). Using Scheffé's technique, it was found that the average of the two experimental groups was significantly better than the control group at the .05 level ($p < .05$); also, the Preston-Botel method was significantly better ($p < .05$) than the compressed speech method. The factor of rate operated effectively as a control variable ($p < .01$) to isolate a large part of the error variance. There were no significant interactions.

In the Gates Reading Survey immediate posttest in the subtest of Vocabulary, significance was found for the main effect of sex ($p < .05$) with girls performing better than boys. Marginal significance was found in rate ($.05 < p < .10$).

In the subtest for Speed, significance was found in the factor of rate ($p < .01$) which again operated effectively as a control variable. The other main effects and interactions were insignificant.

In the subtest for Comprehension, significance was found for the main effect of rate ($p < .01$); no other main effects or interactions were significant.

In the Landmark detailed posttest given 6 weeks later, the only significance found was in the control variable of rate ($p < .01$).

Gain scores for the Landmark test and the Gates test are under analyses and incomplete at this time. It is hoped that significance in methods will be found.

Implications

Although the results of this study are incomplete at this time, it appears that the medium of compressed speech may have some value in helping children increase their rate of reading; however, some cautions need to be made. To compress speech on tapes requires having a speech compressor or access to one. For individual districts to purchase a compressor is extremely costly--$3,900. Tape recorders are necessary and listening units are needed if material is going to be individualized. A tape recorder such as the Wollensak model T1520, which was used in this study, sells for $159.60, and the accompanying listening unit, model HB-4, for $59.95. There is a national center for rate-controlled recordings located at the University of Louisville, Louisville, Kentucky. The Center does provide assistance at a minimal cost in the preparation of time-compressed or expanded recorded tapes of speech for use in experiments and demonstrations. In the future it is anticipated that other centers will be developed where recordings can be made, perhaps at the university level. It appears that in the future districts will have access to a center and having recordings made should not be a difficult problem. An important consideration before sending a tape to a center is the quality of recording on the original tape. This can be extremely critical.

Compressed speech is only one medium to evaluate in considering ways to help children increase their rate of reading. Many approaches are needed. No one mechanical method appears to be superior to others. Even the Preston-Botel method, which is very inexpensive and requires no mechanical gadgetry, proved to be effective and was significantly better than the other two groups.

This was the first study, to the investigator's knowledge, where an attempt was made to individualize the tapes in accordance with the reading rates of the pupils. It is important that more studies be conducted with this in mind if realistic results are to be obtained. The study was limited to sixth-grade gifted children. The results may have been quite different if the Ss had been regular children or retarded children. There is a need for basic research with other kinds of children.

The implications for the use of compressed speech in a public school setting are many. This medium offers great potential not only in the area of developmental reading, but in the areas of remedial reading, speech improvement, listening, and others. It is a matter of availability of the medium, and familiarity and acceptance by the district. Schools today should be exploring and seeking innovative approaches to solving educational problems. Programs per se should not be bought. There are no panaceas. Persons in responsible, decision-making positions should be objectively evaluating and assessing the many kinds of hardware and software now available on the educational market.

CHAPTER XXIV

USING COMPRESSED SPEECH TO TEACH INSTRUCTIONAL TECHNIQUES TO AIR FORCE OFFICERS

Meredith W. Watts, Jr.*

With the dissemination of information concerning the efficacy of accelerated speech as a teaching innovation, more educational institutions may well consider whether the compressed speech format is appropriate for their student clientele. To date, studies have shown that various student populations can adjust to and learn from audio tapes that have been altered to increase the temporal rate of information transfer without substantial loss of comprehension. A general figure of 275 words per minute (wpm) seems to constitute a threshold below which comprehension is not seriously impaired. If these basic studies have been correct, it should be possible to take a relatively standardized curriculum and compress portions of it without suffering a reduction in student achievement. Although various experimental studies suggest that this is the case, few attempts have appeared in the literature to document operational success of compressed speech used in this manner.

The current study reports an attempt undertaken at a military teachers' college to transfer standardized instructional materials from the traditional lecture format to compressed audio tapes and evaluate their comparative efficacy. Several features distinguish these research efforts from many previous efforts. In the first place, the student population is military (United States Air Force) and composed entirely of adults. Secondly, the clientele are virtually all college educated. Thirdly, the materials used in the study were not adapted from textbooks or novels, but rather, were taken from "live" lectures that are integral to an on-going curriculum. As a result, students interacted with the instructional materials in a more realistic fashion than in an isolated experiment. Furthermore, although the students were informed that they were being subjected to a novel method of instruction, it was made clear that the subject matter was "real" and was an integral part of course completion. The general hypothesis tested was that compressed materials would be at least as effective as the traditional format for transmitting information.

---

*Captain Meredith W. Watts, Jr., USAF, is with the Department of Instructional Technology, Academic Instructor and Allied Officer School, Maxwell Air Force Base, Alabama 36112.

Method

Subjects

Subjects in both experiments were students at a teacher training course conducted at Maxwell Air Force Base in Montgomery, Alabama. All were officers ranging in rank from captain to colonel and all were designated for assignments as instructors or instructor supervisors in the Air Force. There was no significant variation in years of formal education among groups in the experiment--virtually all were college educated. Subjects were randomized as to rank so that there was an equal spread across all experimental groups. Pretesting indicated that entry-level knowledge was statistically equivalent (there were no significant group differences in precourse examination scores as indicated by _t_ tests).

Although two separate sets of experiments were conducted, _S_s were of virtually identical description. All were in the summer classes of the Academic Instructor Course (AIC) and were designated for duty as AFROTC instructors. Precautions taken to ensure preexperiment randomization seemed to be effective.

Apparatus and Materials

In Experiment I the effectiveness of compressed tapes and traditional lectures was evaluated; in Experiment II slides and handouts were added to the compressed tapes to determine whether the increased visualization enhanced learning. The description immediately following is generally valid for both sets of experiments. Variations are treated in more detail in the discussion of findings for each experiment.

The raw materials submitted to rate acceleration were lectures normally performed by instructors at AIC. Lectures were taped and transcribed into working scripts. A skilled speaker with a background in English education performed minor editing on the scripts and recorded them in a studio facility at AIC. The tapes were virtual replicas of the lectures although some modifications were necessary where speaker behavior or use of visuals could not be directly reproduced.

Each of the tapes was then compressed on the Eltro Information Rate Changer to 1.5 times the normal presentation rate.* Although estimates of this sort are approximate, it was judged that the finished tapes contained verbal messages at roughly 240 to 260 wpm. Final tapes were dubbed to 7 1/2 ips for playback on Wollensak reel-to-reel recorders. Students heard the tapes in seminar groupings of eight men each. Earphones for each S were not available, but audio quality was not sufficiently denigrated to impair comprehension. It is suspected, though, that student attentiveness may not have been as great as would have been the case if the experience had been more "privatized" using earphones and study carrels to reduce distractions.

Experiment I

Procedure

Subjects were divided into eight-man groupings and subjected to compressed lectures during that period of the school day when the remainder of the students were receiving the identical instruction in the auditorium from a lecturer. Two groups received compressed lectures while two similar groups in the auditorium were identified as control groups. One seminar in the experimental and one in the control group were given a pretest on the material; all four groups were administered posttests on the subject matter. To insure that test results were comparable, exams were constructed from the script and tapes and an O attended the regular lecture to check test items against the "live" performance. Lecturers in all cases adhered to their lesson plans and the accomplishment of course objectives in both lectures and tapes were judged to be comparable.

Subject matter included in Experiment I included three lectures dealing with teacher training. They dealt with three teaching methods ordinarily presented to officer classes at AIC--the Teaching Interview, the Guided Discussion, and the Lecture.

Results and Discussion

In each case the research hypothesis stated that compressed speech would result in greater student achievement than the traditional lecture. The null hypothesis was that there would be no difference in the efficacy of the two modes of presentation. In pragmatic terms, acceptance of either the research or the null hypotheses would be considered a validation of compressed speech, since a demonstration of equal effectiveness would allow the selection of compressed speech on the grounds of efficiency without jeopardizing curriculum achievement. In other words, a validation of rate-accelerated curriculum materials would be accomplished if the tapes proved to be as good or better in producing student achievement of cognitive objectives as measured by objective tests.

---

*The author would like to acknowledge the assistance of Dr. Herbert Friedman of American Institutes for Research (AIR) who provided much valuable advice during the production of tapes and facilitated access to AIR's speech compressor on which all tapes were processed. As is usual in these circumstances, the author absolves him from any blame for infelicities in our research.

Table 24.1 exhibits the means and standard deviations of the four groups subjected to analysis.

TABLE 24.1

MEANS AND STANDARD DEVIATIONS FOR TREATMENT GROUPS, THE TEACHING INTERVIEW METHOD

Group	Means	S.D.
Standard Lecture (Pretest and Posttest)	10.88	1.36
Standard Lecture (Posttest only)	10.88	1.69
Compressed Lecture (Pretest and Posttest)	13.00	2.06
Compressed Lecture (Posttest only)	12.38	1.58

Table 24.2 presents the results of the analysis of variance for the Teaching Interview. In this, as in the two tests to follow, a 2 x 2 design was used to test for treatment effects and for pretest effects. As the F tests indicate, there is a significant treatment effect but no significant effect due either to testing procedure or interaction. Within the limits of this analysis, compressed speech proves to be more effective than the traditional lecture.

The next two tests of the hypothesis are essentially replications of the Teaching Interview analysis. The one conducted for the Guided Discussion teaching method used identical groups for all treatments and was analyzed in a like manner. Examination of Table 24.3, however, will show that the results were somewhat different. (For brevity, only the summary tables will be shown.)

TABLE 24.2

TWO-WAY ANALYSIS OF VARIANCE, THE TEACHING INTERVIEW METHOD

Source	SS	df	MS	F*
Treatment	26.38	1	26.38	8.03
Test Procedure	.78	1	.78	.24
Interaction	.84	1	.84	.27
Error	91.63	28	3.27	
Totals	119.63	31	3.86	

* F.05 = 4.20
F.01 = 7.64

TABLE 24.3

TWO-WAY ANALYSIS OF VARIANCE, THE GUIDED DISCUSSION METHOD

Source	SS	df	MS	F*
Treatment	2.53	1	2.53	1.27
Test Procedure	5.28	1	5.28	2.65
Interaction	1.53	1	1.53	.77
Error	55.88	28	2.00	
Totals	65.22	32	2.04	

* F.05 = 4.20
F.01 = 7.64

The analysis using the Guided Discussion method did not yield significant differences like those for the Teaching Interview. There were no significant differences in the results from either treatments, test procedures, or interaction. The null hypothesis of no difference must be accepted. It is interesting to note, though, that acceptance of the null hypothesis is in a sense a "moral victory" for compressed speech since it indicates that the compressed format was no worse than the traditional lecture; while objectives were accomplished in much less time.

The third teaching method for which comparisons were made was the Lecture. The only difference in the administration of this analysis is that experimental and control groups were switched; that is, those seminars that had received compressed in the first two analyses were returned to the auditorium, the other two were sent to the seminar for compressed materials.

The switch was made to help check against the limitations of the 2 x 2 design. If the results were in some sense a product of one group's enthusiasm for the experiment or lack thereof, although there was no such indication, the alteration in procedure should provide additional insurance against such artifacts. Despite the reversal of treatments among groups, the results are much like those in the Teaching Interview analysis. Significance is shown for the treatment effect, but not for test procedure or interaction (Table 24.4). The difference, again, is in the direction of compressed speech (the table of means is omitted for brevity).

TABLE 24.4

ANALYSIS OF VARIANCE, THE LECTURE METHOD

Source	SS	df	MS	F*
Treatment	9.03	1	9.03	6.26
Test Procedure	3.78	1	3.78	2.62
Interaction	1.53	1	1.53	1.06
Error	40.38	28	1.44	
Totals	54.72	31	1.77	

* F.05 = 4.20
F.01 = 7.64

The level of significance for the treatment effect is lower than in the first analysis--it reaches only the .05 level of significance. But it is nevertheless meaningful that compressed speech materials fare well against the traditional lecture and show student achievement gains equal to or somewhat greater than the traditional presentation. The three analyses presented in Experiment I demonstrate that, if transfer of cognitive-style information is the goal of the presentation, nothing is lost and time and convenience may be gained by adopting the compressed speech tape mode.

The next question to be asked is whether the bare audio tape format of the compressed speech material can be augmented or improved in some way that can improve comprehension and make it clearly superior to the "live" lecture method for teaching of instructional objectives.* A second series of analyses was conducted to determine whether the addition of visual material could enhance learning. These analyses are reported in Experiment II.

Experiment II

Experiments I and II differed in the time at which they were conducted and in the individual students who took part. Otherwise, materials, apparatus, and general characteristics of the Ss remain the same. The variation in procedure was in the addition of visual stimuli in the form of slides and printed worksheets.

Procedure

Experiment II was conducted in two phases, each phase making use of lectures whose subject matter was different from those in Experiment I. The first analysis used a lecture called Introduction to Evaluation. Students were randomly assigned to six treatment groups of eight each. Two groups served as controls, the remaining four received either a compressed lecture or a compressed lecture supplemented by a set of slides. The slides were not particularly rich in subject matter, but were designed to maintain visual interest and help guard against other visual distractions that might divert student attention from the lesson. Essentially, the visuals were adapted from the live lecture and in no way detracted from the comparability of results. As in previous analyses, tests were constructed using the tapes, then validated against the live lecture to ensure uniformity.

In the second phase of the analysis, a lecture in Lesson Planning was presented to two groups (of 16 each) in compressed format. One group received the tape only, while the other group was allowed to use a printed handout which stressed lesson objectives and encouraged them to take brief notes if they wished. Another group of 16 students was used for control purposes. In this analysis, simple t tests between means were used to test for significant differences due to variations in treatment.

---

*It should be remembered throughout that certain aspects of the live presentation are desirable and are lost through the medium of audio tape. There is much "choreography" and interpersonal contact in the lecture that cannot be duplicated. If the purpose of the lecture is to develop favorable attitudes toward the course, or to demonstrate good platform techniques (as a latent and unstated objective), then the live lecture cannot be supplanted. The comments in this report are confined to the achievement of objectives as tested by objective tests taken by the Ss in both experimental and control groups.

Results and Discussion

Table 24.5 presents the one-way analysis of variance for posttest scores in the first phase of the experiment. As can easily be seen, the miniscule F ratio indicates that there were no significant effects due to the experimental treatments. As a cross check, a t test was run on the means of the highest and lowest groups in the experiment and the results were again negative. The addition of visual materials in the form of slides did not increase learning in any measurable way. It is possible that students felt more comfortable with something to look at (and many expressed this feeling informally to the E); however, no such difference showed up in cognitive achievement of students. While it still seems reasonable that visuals can heighten interest and promote learning, the visuals used here did not result in any measurable increase.

TABLE 24.5

ONE-WAY ANALYSIS OF VARIANCE, INTRODUCTION TO EVALUATION

Source	SS	df	MS	F
Between Groups	8.98	5	1.80	.71
Residual	104.13	41	2.54	
Total	113.11	46*		

*A student was absent from one of the groups at the time of testing but this slight variation in group size was not considered to be important in the light of the small F value. Under ideal conditions, of course, the groups should be of identical size.

The second phase of the analysis made use of handout materials and allowed students to take notes. Students did not seem to attempt extensive notes, but it appeared that they were somewhat more comfortable in being able to engage in the standard culturally-prescribed activity of note-taking (regardless of its value). Again, informal observations and unsystematic probing of student reactions does not directly test the research hypothesis, but it does give some insight into the impact visuals can have on such a student population. (An analysis of student attitudes is currently in preparation.)

The means and standard deviations of the three treatment groups are given in Table 24.6.

TABLE 24.6

MEANS AND STANDARD DEVIATIONS FOR THREE TREATMENT GROUPS, LESSON PLANNING LECTURE

Treatment	Means	S.D.	N
Standard Lecture	8.06	2.30	16
Compressed Lecture	8.47	2.19	15
Compressed Lecture with handout	10.06	1.82	16

T tests conducted between groups one (standard lecture) and two (compressed lecture) showed no significant differences. This finding is in accord with the findings in Experiment I in which we found either no difference between treatments, or a slight difference in favor of compressed speech. T tests run between group three (compressed speech plus handout) and the other groups were significant in both cases. A summary of these tests appears in Table 24.7.

TABLE 24.7

SUMMARY OF T TESTS BETWEEN TREATMENT GROUPS, "LESSON PLANNING" LECTURE

Groups	df	t value	Significance Level
1-2	29	.48	--
2-3	30	2.14	.05
1-3	29	2.64	.01

The results of the second analysis diverge from those in the first analysis. Both tested the influence of visuals but in different ways. The first included slides while the second used printed handouts which encouraged the student to interact with the lecture materials at least in a minimal way. Our findings suggest that the latter technique was the more successful. However, it was known prior to the experiment that the slides chosen for the Introduction to Evaluation did not carry any specific content that would necessarily strengthen the original stimulus. For this reason it would be incorrect to conclude that slides themselves do not add to a compressed speech presentation. As Jester and Travers (1966) found, verbal information can significantly increase the achievement of cognitive objectives. Our more "picturesque" visuals did not measurably enhance learning, though there are indications that they had a beneficial effect on student attitudes toward compressed speech presentations.

In the second analysis it was found that handouts could produce increased learning when combined with compressed speech. Naturally this is not to say that any handout would have such an effect. But it does point to the possibility that learning from compressed speech can be augmented and made more attractive to students by integrating printed materials with compressed tapes.

General Conclusions

The first hypothesis to be tested stated that students would learn as well or better with compressed speech format than with the standard lecture presentation. As it turned out, either no difference between methods was discovered or a measurable difference in favor of compressed speech was found. The original concern of this research is therefore satisfied: compressed speech can be substituted for standard lecture presentations when the criterion for student achievement is attainment of a series of cognitive objectives. That is, transfer of information can be effected as well with compressed speech as with traditional lectures, at least where moderate compression rates are employed and the materials are of no more than medium difficulty.

For the educational decision-maker these findings are of particular interest because they were reached in a relatively "natural" surrounding using adults who were well educated, but had no prior experience with rate-controlled recordings. With such an audience, compressed materials were substituted for instructional materials without degradation of the on-going curriculum for those students. Therefore, compressed speech can be seen as an alternative teaching strategy. If the instructor wished to provide variety in the curriculum, or if he wishes to make materials available for makeup or review, compressed tapes seem not only realistic but potentially very productive. Educators concerned with adult and "continuing" education might find numerous practical uses for compressed lectures, particularly if they could be made part of correspondence courses or simply issued to students for playback in a learning center or on a home recorder.

The second set of experiments attempted to assess the increase in educational attainment of students using compressed speech plus some sort of visual stimulus. There was reason to suspect that students would feel more secure, or less distracted by having something to watch while they listened. There was a strong feeling that the multidirectional attention of students should be controlled with appropriate and productive stimuli.

Findings were mixed. The slides did not improve studen comprehension of materials; however, this may have been an artifact of the visuals employed rather than in the concept of visualization itself. It is still held that appropriate visualization will produce gains in student learning and motivation that will be highly desirable when dealing with mature, adult populations.

The use of compressed speech with handout materials did result in increased student achievement on the multiple-choice tests administered. The handouts were not designed to "teach the test," but rather to provide an outline of major objectives and also provide a format within which the student could make simple notes to aid his recall. Previous experience with compressed speech tapes had indicated that students tended to react to visual distractions and lament the lack of note-taking time when listening to compressed speech. Even though the subject matter was of only moderate difficulty, there appears to be a great dedication on the part of educated adults to the practice of taking lecture notes. It is possible that this tendency is not general in all student populations of comparable level, but it was clear from students' comments after the experiment that they felt that they had achieved more and had a retainable record for future study. Within the bounds of the data reported above, it can be said that printed handouts did aid comprehension and that their integration into "operational" instructional materials is a matter well worth considering for those who would like to use compressed speech in a "real" setting where student achievement of objectives is important for academic as well as experimental reasons. The findings of this set of experiments seem to point strongly to the increased use and development of compressed speech in on-going curricula. Furthermore, the observed achievement of Ss as well as their attitudes toward experimental materials seem to offer strong encouragement for the increased use of compressed speech materials for adult student populations.

CHAPTER XXV

ECONOMIC ANALYSIS OF TIME-COMPRESSED SPEECH FOR INSTRUCTIONAL BROADCAST SATELLITES: A PROPOSAL FOR BRAZIL*

D. T. Jamison**

Introduction

My purpose in this paper is rather different from the purposes of most of the papers you have heard at this conference. Most of those were concerned with reports of research or developments on techniques of time-compressed speech. I shall instead try to look at the factors that influence a decision-maker's choice concerning whether or not to use time-compressed speech in a large-scale educational system. That is, I shall be taking the viewpoint of my profession--economics--in order to try to assess the costs and the benefits of introducing such a technique in a large-scale way. In particular, I wish to look at the potential use of time-compressed speech for the Brazilian instructional broadcast satellite program known as Project Satélite Avançado de Comunicações Interdisciplinares (SACI). The government of Brazil is at present spending rather large amounts of money on design of what an operational system would look like; if there is a go-ahead decision within the Brazilian government, the SACI system should become operational in the middle to late 1970's.

While there has been a little work done on the use of time-compressed speech as an alternative to lectures--rather than in the more usual research which used time-compressed speech as an alternative to reading--the level of effort on this has been insufficient thus far to justify a decision for wide-scale implementation in Brazil. There are a number of reasons for this. First, the technique has rarely been used over protracted periods of time. Second, while it appears probable that there will be no problem in using time-compressed speech in Portuguese, this does remain to be tested. Third, the satellite system is a broadcast system, and it would be important to test time-compressed speech in broadcast context rather than in a tape-recorded context if for no other reason than to ensure that it will indeed work well in such circumstances. Finally, as mentioned before, previous research has been primarily oriented toward viewing time-compressed speech as an alternative to reading, and we must test how well this technique holds up in a lecture context.

---

*The author is indebted to Dr. Emerson Foulke for a helpful conversation concerning aspects of the experimental design for the experiment proposed here.

**Dr. D. T. Jamison is Assistant Professor, Department of Economics, and Member, Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford, California 94305.

For the reasons sketched in the preceding paragraph, then, I do not feel that there is at present sufficient evidence upon which to base a decision; therefore, it seems to me that the appropriate thing to do is to gather more evidence in an actual Brazilian setting. It is fortunate that the Brazilians plan in the early 1970's to have a trial run of their broadcast satellite concept, probably utilizing the National Aeronautics and Space Administration's (NASA) Applications Technology Satellite (ATS), series F or G. The ATS-F or ATS-G would be launched around 1972 or 1973 and, if the Brazilians do have access to time on this satellite, would provide a test-bed in a region of northeastern Brazil for large-scale broadcast experiments in television, radio, and I shall propose, radio using time-compressed techniques. Thus my purpose in this paper is, first, to sketch the considerations involved in deciding whether or not the Brazilians should proceed to seriously analyze the possibility of using time-compressed speech in an operational system; and, second, to suggest that the opportunity for experiment posed by the ATS-F or ATS-G be utilized to experiment with time-compressed speech. To provide background I will, in the next two sections of this paper, provide an overview of the Brazilian Project SACI, then describe some of the results that have been obtained so far in the use of time-compressed speech. In the section entitled "Potential Benefits from Using Compressed Speech," I will sketch some of the possible benefits to be expected from utilization of time-compressed speech; in the section entitled "Costs of Introducing Compressed Speech on Project SACI," I will look at some of the costs. Finally, in the section entitled "A Proposed Experiment for the Applications Technology Satellite," I propose an experiment, in a very preliminary way, that could be used to help decide whether this would be applicable for Project SACI.

An Overview of the Brazilian Broadcast Satellite Project SACI

Let me begin by briefly describing a communication satellite system. The communication satellite acts, essentially, as a broadcast repeater.

Instead of having the repeater or transmitter be situated at the top of a tower, the transmitter is located within the satellite. The satellite receives a radio or television signal beamed up from an earth station on the ground and then beams it back to the earth after amplifying it and focusing it. The advantage gained by using a satellite is simply that a region of continental or hemispheric size can be reached with a single communications relay.

In order to have this large-scale coverage be continually available, it is necessary that the satellite be placed into what is called a stationary orbit. A stationary orbit is such that the satellite is in the plane of the earth's equator and revolves in a circle around the center of the earth in exactly the same period of time that it takes the earth to rotate on its axis. A satellite in such an orbit will appear to be fixed in the sky from any point on the earth's surface from which it is visible; hence, the term "stationary." The height of such an orbit is at about 6 earth radii or slightly over 20,000 miles from the surface of the earth.

We might distinguish among a number of different types of communication satellites of this stationary sort. The first and, at present, only existing satellites of this type are point-to-point relay satellites. These satellites have two general characteristics. First, they have relatively low electrical power and hence cannot broadcast a powerful signal. Second, their antennas are quite unfocused so that the beam is spread over the oceans and space as well as over the usage areas for which it is intended. Thus it requires very complex and costly ground equipment to transmit and receive from such satellites, and their use is limited to relatively high-density trunk lines such as over the North Atlantic or from the west coast of the United States to Asia. On such lines relay satellites are, however, extremely profitable. A second class of satellite may be called a distribution satellite. These satellites would be somewhat more powerful than relay satellites and have somewhat more focused beams. This would enable their signals to be received by rather less expensive ground stations than required for those of relay satellites. The television networks in the United States are continually pressing to have such a distribution system established here in the United States; the networks would send up their programs to the satellite from a single relatively expensive ground station and have it received by each of the individual affiliate TV stations across the country on the down link. The third class of communication satellites might be called the broadcast satellites. These are satellites sufficiently powerful and focused in their beam so that ordinary television or radio sets, perhaps upgraded with several hundred dollars worth of additional receiving equipment, can receive their signals. (The term "broadcast" is occasionally used to designate a type of satellite that would transmit directly to an unmodified home receiver. I am in favor of the alternative usage defined here because broadcast satellites of that sort are technologically and economically unfeasible for probably at least 15 years.) In May 1968 the Brazilian National Space Commission, under the leadership of Dr. Fernando de Mendonça, published a three-volume analysis of how broadcast-type communication satellites could be used to rapidly and dramatically upgrade the level of educational opportunity throughout all Brazil. The reason the authors of this report considered it so essential that this be done was the very high illiteracy rate existing throughout all of Brazil and, in particular, in its vast and relatively undensely populated interior. The unique capabilities of a communication satellite to reach an area as huge as Brazil made the satellite concept appear extremely attractive for the instructional purposes they had in mind. In addition this report proposed that the communications satellite be used to upgrade the conventional telecommunications network by providing both telephone and telegraph services throughout much of the interior of Brazil. At present the government of Brazil is spending, through its National Space Commission, on the order of a million dollars a year for further analysis and design studies of this concept. They have earmarked large sums of money for an experiment in the northeast of Brazil utilizing NASA's ATS-F or ATS-G experimental communication satellite mentioned in the Introduction. They have submitted for this purpose to NASA a formal proposal for time on one of those satellites. It should be noted that not only is Brazil studying carefully and investing in this concept, but the government of India, faced with very similar problems though on an even larger scale, is also working with this idea. Sometime previously to the government of Brazil, they had submitted to NASA an application for an experiment on the ATS-F, and late in the summer of 1969 an agreement was signed between the governments of the United States and India for cooperation on such an experiment. Most of the comments that I have to say in the rest of this paper concerning Brazil apply equally well to India.

In the May 1968 report of the Brazilian government, the medium suggested for instructional use was television. In previous work for the Brazilian National Space Commission and in collaboration with a number of their employees, I have coauthored several papers suggesting an alternative to this approach--Jamison, Ferraz, and Torquato (1969) and Jamison, Jamison, and Hewlett (1969). In these papers evidence was reviewed concerning the relative instructional effectiveness of television and radio and suggesting that radio is almost as good or equally as good as television for most instructional purposes. In addition the cost of radio in both dollar and other terms is much less than that of television. The conclusion seemed clear that in addition to television, or perhaps instead of it, a large number of radio channels should be used as the primary means of instructional communication. The thrust of the two papers cited was to describe a number of alternative uses for such a relatively large number of radio chann[illegible]. It may well be that the most cost-effective way to use these radio channels for instructional communication is to utilize them with time-compressed speech. It is my purpose in this paper to analyze that possibility.

Experiments with Time-Compressed Speech

There have been a large number of ingenious experiments concerning how to best produce and utilize time-compressed speech. These experiments have involved a rather detailed consideration of a number of the factors that are conducive to effective utilization of time-compressed speech and those factors that are not. I cannot in a few paragraphs hope to do any justice to this literature and, fortunately, there is no need to do so. A recent paper in the Psychological Bulletin by Foulke and Sticht (1969) provides an excellent survey of this literature. I shall just state briefly some of the results they report.

First, there is the question of what ways may be used to increase the rate at which speech is produced. That is, how do you take an ordinary tape of a human lecture or reading session and produce a tape that can be played back at much higher speeds and still be clearly understood. The first and most obvious way is simply to speed the tape up. Unfortunately, this causes the same sort of frequency distortion that occurs when you play a 33 rpm record at 45 rpm. The most common way to speed up the rate of delivery of speech that is now in common use involves, essentially, the removal of 10 millisecond (msec.) chunks of the tape every so often and then jamming the rest together. The frequency of the "every so often" determines the degree of compression. A good deal of research has been expended on how this might be done with particular reference to the problem of how the ends of the remaining sections of speech may be abutted in a way that sounds proper. A third and more expensive approach to compression of speech involves computer transformation of the tape. This appears quite promising for high quality compressions in the future though it is very expensive if the tape is to be heard by only a few people.

A second class of investigations on time-compressed speech has centered around the question of how comprehension is affected by increasing the effective rate of delivery of speech. Normal human speech may be considered to have a rate of delivery of approximately 175 words per minute (wpm) though there is considerable variation among speakers. Two sorts of experiments have been run to see how the intelligibility of compressing speech to a rate of delivery considerably higher than 175 wpm is affected. In the first, single words were compressed and the subject is asked, after listening to the compressed word, to repeat it. In the second type of experiment, large blocks of connected discourse are presented to the subject by way of tape and he is asked, at the end of an experimental session that might last many hours spread over a number of days, to answer questions on a test that is to measure his mastery of the subject matter. It turns out that large fractions of a word may be discarded in the single word mode and have the subject still be very capable of understanding what that word was. However, this is of little interest to our present application. For connected discourse when the word rate reaches 275 wpm or more, it does appear that there is a significant degradation in comprehension. It appears from this research that around 250 wpm might be best for a lecture usage. Further research may be required on this point, however, depending on the exact context of usage.

A third sort of investigation on time-compressed speech concerns how training in listening to time-compressed speech affects a subject's capability to comprehend at relatively high rates of delivery. Most of the results reported thus far in this area are disappointing. Training does not appear to be capable of significantly improving students' comprehension of compressed speech.

Finally, it is worth making a few comments on the general observation that students have a positive attitude toward utilization of time-compressed speech. This has been the general observation of the experimenters in this area; see, in particular, the comments of Friedman and Orr (1967). They assert, for example, that: "An overwhelming number of experimental subjects have had a favorable attitude toward time-compressed speech [p. 69]."

Potential Benefits from Using Compressed Speech

There are two broad classes of benefits that may be expected from utilizing compressed speech in a large-scale operational system. The first class of benefits is concerned with more efficient utilization of the time of the students within the system, and the second class of benefits stems from a more efficient utilization of the expensive system hardware.

One of the largest costs to an economy of having students in school is that of the students' time. If the student were not in school, he could be constructively employed in the economy; and, therefore, an opportunity cost to the society of having him in school is the benefit of the labor he could have performed on the outside. The magnitude of this cost has been the subject of a number of recent economic investigations that are surveyed by T. Schultz (1963, pp. 27-32). The following result seems to hold from approximately the junior high school level to the university level, independently of the status of the economic development of the country: approximately one-half the cost of education is that of the students' time. Another way of putting this is that the value of the earnings foregone by students while they are in school is approximately equal to the total cost of providing classrooms and teachers to the economy. If we assume that the student can understand the material just as well if it is presented to him at a rate of delivery of 250 wpm as he can at a normal rate of delivery, then it will only take him approximately seven-tenths as much time to go through the same amount of aurally presented material. This timesaving is a direct economic benefit, and it is possible that specific work can estimate both the magnitude of the timesaving and its actual value as a function of the subject matter and educational level of the student. This is one area of work that deserves a good deal of further attention.

The second broad class of benefits from using time-compressed speech is associated with that of saving time on the expensive hardware that would have to be constructed for a system such as SACI. The level of the investment that we are talking about here is on the order of several hundred million dollars--on the order of fifty to a hundred million dollars for the satellite system and probably a thousand dollars each for several hundred thousand or more ground stations. The exact numbers here depend very much on the actual system design chosen, but these numbers give a ballpark estimate of the really quite large quantities of money that are involved. It is thus a matter of some interest if it is possible to save 10% on the utilization of the system. If the same bulk of instructional material can be presented in 90% of the time that it would otherwise have taken, then either one of two things could be done. First, the entire system could be designed in such a way that this timesaving could be used to reduce the overall costs of the system. Or, second, the time free from instructional utilization could be used to increase the volume of telecommunications services made available by the satellite and hence the revenue generated from such services.

Costs of Introducing Compressed Speech on Project SACI

In the preceding section we outlined a number of potential benefits from introducing compressed speech on SACI. In this section I would like to look at the other side of the picture. There seems to me to be a number of potential cost areas that should be considered and these will be taken up in the paragraphs that follow.

The first cost area introduced by compressed speech would be that of actually preparing the tapes with compressed speech from those given at an ordinary level of delivery. This cost is, I feel, one that we can safely neglect in the computations. The reason is simply that the cost of preparation of a tape is quite small on a per capita base when the number of users of the tape is in the hundreds of thousands or perhaps even millions. That this cost may be considered unimportant implies that it is probably reasonable to obtain the highest quality compression possible, including consideration of utilization of computer compression.

A second cost area may be in the cost of obtaining additional bandwidth per radio channel. It may well be, and this I do not know, that even with a quite high signal-to-noise ratio (s/n), time-compressed speech cannot be reproduced with relatively high fidelity in the three kHz that is normally allotted to a telephone channel. Thus it may be necessary to increase this bandwidth allocation somewhat and thus use up additional chunks of the valuable radio spectrum. It is not my present feeling that this cost would be a very large one though it is certainly a point that should be examined further.

A third potential cost area for use of compressed speech involves the high probability that in order to maintain a constant level of comprehension with time compression, a higher level of output s/n may be required. The reason for this is simply that given a relatively high rate of delivery, people are much more likely to lose the thread of what is being discussed with a relatively small disturbance. For this reason signal quality is quite important. Unfortunately, the overall cost of a broadcast satellite system is quite sensitive to the output s/n. For a fixed number of channels and fixed ground receiver sensitivity, the only way to obtain a 3 dB improvement in s/n would be to double the effective radiated power output of the satellite. If the satellite is constrained to broadcast into a fixed geographical region, this means doubling the raw power output at a vast increase in the cost of the satellite not only in terms of its dollar cost but also in terms of its expected reliability. Fortunately, we can obtain a reasonably good estimate of these costs from the type of computer model that is reported, for example, in Haviland (1968). The other input required for analysis of this cost aspect of compressed speech is that of precise specification of how reducing the s/n degrades the performance of the student when listening to compressed speech. Results on this point in the literature need to be cast into a framework amenable for this type of analysis and, perhaps, additional work needs to be done of an experimental nature on this point. My present feeling is that the cost of required improvements in the output s/n will be the most important addition to system costs that including time-compressed speech will generate.

A fourth potential cost area is that with time-compressed speech the students may be required to listen through earphones rather than simply having a loudspeaker broadcast to the class. For example, in the summary concluding the Proceedings of the 1966 Louisville Conference on Time-Compressed Speech (1967), the following is asserted: ". . . many signal distortions which are not critical in the reproduction of speech at normal rates may become critical at accelerated rates. Knowledge of the contributions of various kinds of distortion should be used in stating the design criteria for playback equipment. The choice of earphones or loudspeaker constitutes a simple illustration. It has been found that highly compressed words are significantly more intelligible when heard by means of earphones, instead of a loudspeaker. This is undoubtedly due to damping problems, inherent in loudspeakers that are avoided in earphones [p. 152]." While earphones are certainly desirable to be included in the system at any rate, so that different students in the same classroom may be listening to different programs at the same time, some part of this cost should be attributed to compressed speech if, indeed, it is a requirement for that utilization.

The above cost discussion is meant, clearly, to merely delineate what the sources of additional costs are likely to be, not to specify in detail what those costs are. Only a good deal of more detailed information that will become available later, hopefully, will enable us to specify very clearly what the total cost increments will be. However, it does appear at present that these costs will not be unduly high unless the comprehension of time-compressed speech is terribly sensitive to the s/n.

A Proposed Experiment for the Applications Technology Satellite

I would like to conclude this paper by suggesting that a simple experiment be included among the radio experiments that have already been proposed by the Brazilian government to NASA for inclusion in the ATS-G package of experiments. There would be six different treatments for the subjects in this experiment, and these treatments are illustrated in Table 25.1.

TABLE 25.1

TREATMENTS IN PROPOSED ATS-G BRAZIL EXPERIMENT

		Signal-to-Noise Ratio	
		High	Low
Amount of Compression	None	I	IV
	250 wpm (Compression used to save time)	II	V
	250 wpm (Compression used to present more material in same time)	III	VI

There are two different levels of s/n included in the experiment to provide a fairly clear test of how important this variable is in an operational context. There are three different alternatives presented under rates of delivery. The first would be at a normal rate of delivery and this would serve as a control group. The second would utilize time-compressed speech in order to save time; that is, the students in this group would have exactly the same lecture as that of the control group but would receive it in something like 70% of the time if, say, the rate of compressed delivery were 250 wpm. The third group would not use the compression in order to save time but rather in order to obtain more material. That is, for this group a different set of lectures based on exactly the same reading material would be prepared and the length of those lectures would be such that students listen to lectures for the same period of time as the control group but, clearly, would be receiving a good deal more information. In this case if the rate of delivery were 250 wpm they would be receiving something over 1.4 times as much lecture material.

It is certainly a matter for later decision, and a decision to be made in Brazil, as to what grade level and what subject matter should be chosen for the experiment. I feel that probably most appropriate for an initial selection would be at the university level in a social science or humanities subject matter. It is certainly possible to think of a number of simultaneous experiments being conducted at a number of different grade levels or in a number of different subject matters. The material presented over the air in the experiment would be a complete replacement for the ordinary classroom presentations for the entire duration of, hopefully, a year-long course. If this were to be done, we would obtain a clear idea of the long-term effects of utilizing compressed speech in an ordinary educational environment.

The experiment that has just been sketched differs in a number of important ways from most of the experiments that have already been done in time-compressed speech. First, the experiment involves replacement of a lecture rather than replacement of reading by the compressed speech techniques. Meredith Watts, in a paper in this volume, has reported one instance of utilization of compressed speech for lecture purposes.* It is very encouraging to note that the subjects in his experiment responded quite favorably to this use of compressed speech. A second difference between this and previous research is a small one, but one that should be looked into, and that is that the compressed speech would be broadcast rather than simply recorded. It is important, I think, to gain experience with this way of distributing compressed speech. Third, the language for research on compressed speech has heretofore been almost entirely English. In this experiment the language would be Portuguese and it will be, I think, of some interest to see if the same compression techniques that work for English are also suitable for Portuguese or if some other compression techniques might be required. It might also be that languages differ in the ease with which they may be compressed without loss of comprehension; this will be important to look into. Fourth, as previously mentioned, this would be a year-long experiment in an essentially operational environment for the student--that is, he would be motivated for this course in exactly the same way that he would be motivated for any of the rest of his courses. We will thus provide a means for studying in a very real-life rather than contrived situation what the long-term effects of time-compressed speech might be.

---

*See Chapter XXIV, "Using Compressed Speech to Teach Instructional Techniques to Air Force Officers."

In closing I would like to note my own belief that time-compressed speech appears at present to be worthy of very serious consideration for Project SACI. The costs do not appear to be high and the additional motivating factor of compressed speech plus its capability for more rapid delivery of information give promise that it will be a useful technique for an operational system. An experiment such as the one that I have proposed here could test the validity of these assumptions and provide generally relevant background information to other researchers in this area.

CHAPTER XXVI

AN INVESTIGATION INTO EXTENDED USE OF TIME-COMPRESSED SPEECH WITH INTERMEDIATE-GRADE SUBJECTS*

Grace D. Napier**

Problem

Is time-compressed speech, within the limits of 160 through 367 words per minute (wpm) rates employed in this study, feasible as an adequate avenue for learning certain types of material when intermediate-grade Ss have been provided with extended, systematic training in its use through gradual acceleration of word rate and frequent tests on comprehension?

Method

Experimental Design

Nine compression levels beyond the base rate of 160 wpm (183, 206, 229, 252, 275, 298, 321, 344, and 367 wpm), five types of test items (factual, "because," vocabulary, true-false, and heard-not heard), and placement of three quizzes in each 2-day lesson (immediately after presentation of passage on first day, at beginning of second day to test delayed recall, and immediately after presentation of passage on second day) produced comprehension scores during training.

Two criterion tests measured differences for experimentals before and after training and for controls who had no training in the interim.

---

*The research described in this report was submitted by the author as an abstract of a dissertation submitted to the Temple University Graduate Board in partial fulfillment of the requirements for the degree, Doctor of Education.

**Dr. Grace D. Napier is Assistant Professor of Special Education at the University of Northern Colorado, Greeley, Colorado 80631.

Criterion Instruments

The Cooperative Sequential Tests of Educational Progress: Listening, Form 4b and the Durrell-Sullivan Reading Capacity Test, Intermediate Test A served as pretraining criteria, while Form 4a of the former and the identical form of the latter were posttraining criteria.

Subjects

Sixty-six intermediate-grade pupils (boys and girls) enrolled in Logan Public School, Philadelphia, Pennsylvania, served as experimentals. Fourth- and fifth-grade visually handicapped (blind or partially seeing) Ss and fifth-grade normally seeing Ss constituted experimentals. Controls were drawn from the population of the same school. Chronological age, sex, visual reading score, intelligence quotient, vision classification (blind, partially seeing, or normally seeing), and grade placement were considered as contributing factors.

Experimental Materials

Meet the Presidents by Frances Cavanah and Elizabeth L. Crandall (1962) had been commercially sound-recorded and later time-compressed on a Tempo Regulator. The 34 chapters, the entire book, were divided into nine groups in order to be time-compressed at specific wpm rates progressing from the base rate of 160 to 183 through 367 wpm as indicated above. Experimentals heard each chapter twice; the second playback was 23 wpm faster than the initial presentation of that chapter. Thus, the two presentations of the same chapter became the 2-day lesson plan described earlier.

According to the Dale-Chall Readability Formula, this book--biographical sketches of the United States presidents--has fifth-grade readability.

Five hundred and ten teacher-made questions tested comprehension during training. Each 2-day lesson employed 15 questions from the battery. These 15 were administered in three quizzes of five items each. Each quiz included one each of the five different types of items: factual, "because," vocabulary, true-false, and heard-not heard. The battery served 68 school-day training sessions.

The Sequential Tests of Educational Progress (STEP): Listening, Form 4b was administered unmodified, with the E reading the test aloud to the Ss according to instructions in the accompanying publisher's manual. Form 4a was modified (with the publisher's permission) so that the passages usually read live to the Ss were sound-recorded and then time-compressed to playback at 300 wpm. The presentation of 4a was the first and only time that control Ss experienced time-compressed speech.

The Durrell-Sullivan (with only one form available) was presented as pretraining criterion and repeated unmodified as posttraining criterion. Although this is a picture-type test, partially seeing Ss had no difficulty working with these pictures, though blind Ss, of course, did not participate in this testing phase.

Response cards were used during the training period. Mechanics involved in using response cards were kept to a minimum in order not to contaminate results; i.e., since the experiment was investigating listening comprehension, results must not be contaminated by problems of spelling, penmanship, etc. All that Ss had to do was write their names on the cards and check the appropriate box for each answer to multiple-choice questions (see illustration, Appendix A).

Procedure

Two criterion instruments were administered to experimentals and controls before the training period. Only experimentals were, during training, exposed to regulated increments of time-compressed speech and tested on comprehension. After training, two criterion tests were administered to both experimentals and controls.

Hypotheses

Eight null hypotheses were tested for statistical significance. All hypotheses except Hypothesis 5 were rejected at the .01 level of confidence; Hypothesis 5 was rejected at the .05 level of confidence.

H_1: there is no significant difference for quiz scores from one compression level to another among the nine employed (160-183 wpm, 183-206 wpm, 206-229 wpm, 229-252 wpm, 252-275 wpm, 275-298 wpm, 298-321 wpm, 321-344 wpm, 344-367 wpm).

H_2: there is no significant difference for scores between fourth- and fifth-grade Ss being trained.

H_3: there is no significant difference in quiz scores between and among the five test items (factual, "because," vocabulary, true-false, heard-not heard).

H_4: there is no significant difference for quiz scores among the three quizzes administered during each 2-day lesson.

H_5: there is no significant difference in test scores attributable to difference in intelligence quotients.

H_6: there is no significant difference in test scores attributable to difference in chronological age.

H_7: there is no cutoff point on the pretraining tests that could serve as a predictive measure for success in training with time-compressed speech.

H_8: control Ss do not demonstrate difference between pretraining and posttraining tests comparable with the difference between tests for experimental Ss.

Limitations

The study was characterized by the following limitations observed in the experiment:

1. The experimental design was pretests, training with measurement of performance, and posttests.
2. The experiment was based on listening and evaluation of listening comprehension.
3. The experiment was scheduled as but one part of the school day for more than 4 months with the routine problems of interrupted schedules, extraneous noise, absenteeism, discipline, and differences in readiness or receptivity, ability to learn, degree of motivation, etc.
4. Material listened to was of only one type, namely, biographical sketches of the American presidents.
5. Each biographical sketch was limited to a 2-day lesson plan regardless of the length of the sketch or its historical significance.
6. Test items during the training period were of only five types and, therefore, do not account for all types possible.
7. Experimental Ss were limited to fourth- and fifth-grade placement in the elementary school.
8. The mechanics of the quizzes during the training period were easy for the Ss to execute on response cards and required no specific skill except the ability to check the appropriate square and writing one's own name.
9. The experiment was limited to the playing of tape-recorded material and involved little active teaching on the part of the examiner.
10. The criterion instruments were limited to two listening or hearing comprehension standardized tests.
11. This experiment was limited in its use of time-compressed speech to the pattern whereby the experience began with normal word rate and introduced a predetermined increment at regular intervals throughout

the training period.

12. This experiment, in its quiz items, used from two to five answer responses depending on the type of quiz item employed. Therefore, results must be interpreted with this fact in mind.

Results

Fourth grade--though having significantly lower means during training than either fifth grade and in spite of lower grade placement, youngest mean chronological age, lowest mean intelligence quotient, and highest mean absentee rate--showed a more steady upward climb for the nine compression levels, indicating that fourth grade learned more in relation to its base rate than either fifth grade in relation to fifth-grade base rates. Fourth grade, like both fifth-grade groups, earned its lowest mean on Quiz 2 (delayed recall) but was the only group manifesting rise on Quiz 3, indicating that the second presentation of the chapter was beneficial to fourth grade but seemingly not to both fifth grades. Fourth grade, earning its lowest mean on vocabulary (also true for both fifth grades), exhibited greater ability to discriminate between heard and not-heard than between true and false, whereas both fifth-grade groups showed little difference between these two types of test items, perhaps little more than chance guessing in both cases. A word must be inserted here regarding performance on vocabulary. Though all Ss did poorly here during training, performance on the Word Meaning section of the Durrell-Sullivan was considerably better (see Limitation 12).

Criterion Instruments

On the STEP, employing time-compressed speech after training, fourth grade did as well as both fifth grades combined, with normally seeing fifth grade being better and visually handicapped fifth grade poorer. Experimentals excelled controls. Experimental girls and boys did equally well, whereas control boys surpassed control girls.

On the Durrell-Sullivan, Word Meaning, fourth grade achieved greater point gains than either fifth grade. Visually handicapped Ss surpassed the normally seeing. Experimentals excelled controls. In Paragraph Meaning, fourth grade excelled both fifth grades. Visually handicapped did only slightly less well than normally seeing Ss. Controls excelled experimentals in this area.

On the entire Durrell-Sullivan, fourth grade exceeded both fifth grades; visually handicapped surpassed normally seeing Ss. Controls exceeded some experimentals, but controls surpassed none of the fourth-grade experimentals. Experimental boys surpassed experimental girls, while the contrary was true for controls.

Discussion

Eight null hypotheses were tested and rejected. Fourth grade excelled both fifth grades, and often visually handicapped excelled normally seeing Ss. When time-compressed speech was used in testing after training, experimentals surpassed controls. In spite of seeming disadvantages, fourth grade proved that training alone accounted for success. Why the two fifth-grade experimental groups did not also prove this so strikingly is not altogether clear. Evidence suggests that in this study Ss of lower intelligence, lower grade placement, youngest chronological age, and highest absentee rate might have been more teachable when time-compressed speech is employed. This prompts the recommendation that time-compressed speech should be investigated with younger children and those who might be of lower intelligence. One can be extremely optimistic about the educational value of time-compressed speech as a communication medium.

APPENDIX A

RESPONSE CARD

1 A B C D E

2 A B C D E

3 A B C D E

4 A B C D E

5 A B C D E

CHAPTER XXVII

PROGRAM OF TRAINING IN LISTENING-READING FOR VISUALLY IMPAIRED STUDENTS USING COMPRESSED SPEECH RECORDINGS

Rachel Rawls*

This study examined the assumption that reading by listening is a language skill involving certain cognitive processes. These processes are related both to listening in any other type of listening situation and to reading. Also, improvement should result from instruction and practice.

A second assumption was that compressed speech recordings could be employed in the training program effectively, and that reading efficiency using listening would increase both uncompressed and compressed speeds after a program of listening instruction and practice.

Experimental programs of listening-training were administered to high school students at the Governor Morehead School, Raleigh, North Carolina. A fourth group of students served as controlled *S*s. All of the students in the tenth, eleventh, and twelfth grades participated in one of the groups; each was assigned either to a treatment or controlled group on a random basis. The distribution of *S*s appears in Table 27.1.

Before starting the training programs, a battery of tests was administered to all of the participants. These included: Verbal Scale, Wechsler Adult Intelligence Scale (WAIS); Gilmore Oral Reading Test, Form A; Slosson Oral Reading Test (SORT); and Brown-Carlsen Test of Listening Comprehension. A silent reading rate was established by having a 5 minute timed period using *Everyweek*, a weekly news magazine for high school students, and counting the number of words each student read.

---

*Mrs. Rachel Rawls is in the Department of Psychology, School of Education, North Carolina State University at Raleigh, and is Psychological Consultant for the North Carolina State Commission for the Blind, Raleigh, North Carolina.

TABLE 27. 1

DISTRIBUTION OF STUDENTS PARTICIPATING BY GRADE LEVEL AND MODE OF READING

Primary Reading Mode Preference	Grade Level 10	11	12	Total
Braille	14	7	10	31
Print	5	2	7	14
Total	19	9	17	45

The mean IQ for the entire group was 106.008 (S.D. = 13.181). For braille readers the mean was 106.452 (S.D. = 13.363), and for print readers 105.286 (S.D. = 14.346). The difference between braille and print readers was not significant. Differences among the groups were not significant.

On all of the other tests administered, the only significant differences were found on oral and silent reading rates between braille and print readers. Among braille readers and print readers compared by groups, however, there were no significant differences in words read per minute. Table 27. 2 shows the means for braille and print readers in each group on the Gilmore Oral Reading Test and for the silent reading rate.

None of the other tests revealed differences among the groups nor between braille and print readers in these groups. The differences on reading rate would normally be anticipated when comparing low vision print readers with braille readers.

Table 27. 3 shows the means for each of the other pretests, the scores on the WAIS Verbal Scale, SORT, Gilmore Oral Reading Test, and Brown-Carlsen Test of Listening Comprehension.

The distribution of braille and print readers in each of the groups appears in Table 27. 4.

At the conclusion of the listening practice training program which will be described, participants were asked to take several of the tests again to document whether changes did occur as a result of the training. The three scores of the Gilmore Oral Reading Test (accuracy, comprehension, and oral reading rate) were obtained using Form B; three subtests of the Brown-Carlsen Test of Listening Comprehension were repeated: recognizing transitions, word meaning, and lecture comprehension. The SORT was readministered. A silent reading rate using the same method was established at the conclusion of the training programs. In addition Listening Comprehension tests were given six times during the series of listening practice sessions. At the end of the practice sessions, two forms of the Sequential Tests of Educational Progress (STEP) Listening Test, Level 2, were given. One of these was at 30% compression and the other at regular recorded speed.

Listening Reading Training Programs

Two approaches to listening-training were adopted. Previous studies of both visually impaired and normally sighted groups had yielded somewhat equivocal evidence about the probable benefit of practice, when using compressed speech recorded materials, in increasing capacity to comprehend while listening. One group, therefore, had 16 listening practice sessions, each followed by a short list of comprehension questions. The sessions began with a story heard at the regular recorded speed. Subsequent practice sessions were heard at 10%, 20%, 30%, and 50% compression.

A high school textbook, Perspectives, was recorded for the practice sessions. This book is available in a volunteer produced braille edition. None of the students had had previous experience with the material in the text.

The second group heard the same materials at the same rates of compression followed by the comprehension questions, but these participants had copies of the text in braille or print before them as they listened. They were requested to try to read this copy while they heard the recording.

The third group engaged in the same practice sessions with the other students. Prior to beginning these, this group had a series of five lessons on listening-reading techniques. These lessons were adapted from Listen and Read, M-P, published by Educational Developmental Laboratories. Modifications were made in the material to make it applicable for the purpose of the study.

Five specific hypotheses concerned with the expected effectiveness of each approach to improvement to improving listening-reading and three hypotheses concerned with possible subsidiary benefits on tactual or visual leading were examined. These hypotheses are shown on later pages. To summarize briefly, the predictions were:

1. A program of lessons in listening-reading techniques, followed by practice sessions using graduated rates of compression, would improve student's ability to listen to and comprehend recorded reading materials more than would the other two approaches.

TABLE 27.2

MEAN READING RATE, SILENT READING, AND GILMORE ORAL READING TEST

Test &	Group 1		Group 2		Group 3		Group 4		Combined	
Reading Mode	M	S.D.	M	S.D.	M	S.D.	M	S.D.	M	S.D.
Silent Reading Rate										
Braille readers	98.7	31.4	128.1	43.9	101.1	51.5	102.6	35.6	108.1	43.3
Print readers	165.8	54.2	177.8	30.3	183.4	37.4	149.0	26.6	167.2	42.9
Gilmore Oral Reading Test										
Braille readers	80.2	24.4	89.4	17.8	71.8	32.4	82.9	19.4	81.2	24.2
Print readers	99.6	10.9	101.1	19.4	97.5	20.9	93.9	2.1	98.7	23.9

TABLE 27.3

MEAN SCORES FOR PRETEST MEASURES BY GROUPS AND READING MODE

Test &	Group 1		Group 2		Group 3		Group 4		Combined	
Reading Mode	M	S.D.	M	S.D.	M	S.D.	M	S.D.	M	S.D.
WAIS										
Braille readers	104.0	13.0	107.0	13.3	109.0	15.2	105.3	11.1	106.4	13.4
Print readers	108.2	6.7	106.5	6.4	96.7	21.9	108.5	8.5	105.3	14.3
Combined	105.9	10.7	106.8	11.7	105.6	18.1	105.9	10.8	106.1	13.2

TABLE 27.3--Continued

Test & Reading Mode	Group 1 M	Group 1 S.D.	Group 2 M	Group 2 S.D.	Group 3 M	Group 3 S.D.	Group 4 M	Group 4 S.D.	Combined M	Combined S.D.
SORT										
Braille readers	186.5	12.3	188.1	6.9	195.2	11.0	183.0	11.9	186.6	10.7
Print readers	187.6	14.1	191.5	2.7	178.7	15.0	194.4	1.5	188.8	17.7
Combined	187.0	9.6	189.2	6.0	183.4	12.8	185.1	11.6	186.3	10.5
Brown-Carlsen Test of Listening Comprehension										
Braille readers	42.0	9.9	43.4	11.0	47.5	8.7	43.5	8.3	44.2	9.5
Print readers	48.0	7.3	46.3	9.0	44.0	12.0	48.5	2.1	46.7	16.5
Combined	44.7	9.2	44.3	10.5	45.5	9.8	44.4	7.8	45.0	9.5
Gilmore Oral Reading Test, Accuracy Rate										
Braille readers	85.5	13.3	88.5	4.4	84.5	9.7	83.2	14.0	85.4	11.2
Print readers	88.0	3.6	90.5	1.1	85.3	10.0	91.5	.2	88.6	5.6
Combined	86.6	10.2	89.2	4.1	84.7	9.7	85.3	8.8	86.5	8.7
Gilmore Oral Reading Test, Comprehension Sc.										
Braille readers	44.8	2.8	45.6	2.7	45.6	2.9	45.5	1.3	45.4	2.6
Print readers	46.4	1.0	45.3	1.1	43.3	2.9	45.5	1.5	45.4	2.4
Combined	45.5	2.4	45.5	2.4	45.0	3.1	45.4	1.2	45.4	2.4

TABLE 27.4

NUMBER OF BRAILLE AND PRINT READERS BY GROUPS

Reading Mode	Number of participants Group 1	Group 2	Group 3	Group 4
Braille	6	8	8	9
Print	5	4	3	2
Total	11	12	11	11

2. Using two modalities to read would be more effective in improving listening-reading than practice alone.
3. Practice sessions would result in some improvement in listening-reading comprehension when participants were compared to students in the controlled group.
4. All approaches to increasing listening-reading ability would have subsidiary benefits on visual and tactual reading, but the greatest benefits would accrue to those students who were exposed to braille or print copies of the listening material as they listened.

Results

During the practice sessions, six Reading Comprehension tests were given at intervals. Three of these were furnished by Dr. Emerson Foulke from materials he had used in previous studies.* These materials varied in level of difficulty from seventh-grade to college-level reading. As can be seen in Table 27.5, Group 3, the group that had listening-reading technique lessons prior to practice, consistently had a higher mean percentage of correct responses on the tests.

Primary Hypotheses**

H_1: initial instruction in effective listening techniques, followed by exposure to oral methods compressed up to 50% of original recording time, will result in increased auditory reading efficiency when this is measured in terms of comprehension and retention.

---

*Dr. Emerson Foulke is Director of the Perceptual Alternatives Laboratory, University of Louisville, Louisville, Kentucky 40208.

**See footnote, page 296.

H_2: initial instruction in effective listening techniques, followed by listening to recordings compressed at increasing percentages of time, will constitute a more effective method of training in aural reading than will practice alone in listening to such passages, followed by comprehension questions (no prior instruction).

H_3: initial instruction in effective listening techniques, followed by listening to recordings compressed at increasing percentages of time, will constitute a more effective method of training in aural reading than will practice in listening to such passages while reading the passages simultaneously in print of braille modes, followed by questions (no prior instruction).

H_4: practice alone in listening to passages recorded at increasing percentages of compression, followed by comprehension questions, will increase comprehension in listening to materials at regularly recorded speeds and to those which have been time-compressed.

H_5: systematic practice with auditory reading at increasing rates of compression, when simultaneously accompanied by Ss' reading the same materials in inkprint and braille, will result in more effective comprehension and retention of materials at regularly recorded speeds and those which have been time-compressed than will practice followed by questions alone.

Secondary Hypotheses**

H_1: all training procedures in listening reading will result in gains in silent and oral reading speed of braille or visual materials.

H_2: subjects having prior training in efficient listening will experience greater gains in tactual or visual reading than will Ss having practice in listening only.

H_3: subjects reading simultaneously in two modalities (visual and auditory or tactual and auditory) will experience greater gains in braille or print reading than will Ss having listening practice only and Ss having prior training in efficient listening.

**See footnote, page 296.

TABLE 27.5

MEAN PERCENTAGE OF CORRECT RESPONSES ON LISTENING COMPREHENSION TESTS ADMINISTERED DURING LISTENING PRACTICE SESSIONS

Test	Percent Compression	Mean percent of correct responses by groups			
		1	2	3	4
1	0	57	57	81	
2	0	71	76	81	
3	0	43	43	57	
4	20	67	67	76	52
5	50	38	33	33	
6	50	43	43	52	38

In order to compare the results of the several tests, standard scores were computed. Analysis of variance of the uncompressed tests (see Appendix Table 27.1), the first three, indicated significant effects for group differences and interaction between tests and reading mode as well as group by reading mode by tests interaction. Only the experimental groups had these tests.

No significant effects were noted for tests' differences so participants were compared across tests. There were also no differences between braille and print readers in performance. Comparing the mean scores among the groups, differences between Group 3 and each of the other two groups were significant, but the differences between Group 1 and Group 2 were not. Means are shown in Table 27.6.

**Subsequent references in the text will appear as:

I-A	Instruction + Practice Increases Auditory Reading Efficiency
I-B	Instruction + Practice > Practice Alone
I-C	Instruction - Practice > Practice + Print/ Braille Reading
I-D	Practice Alone Increases Auditory Reading Efficiency
I-E	Practice + Print/ Braille Reading > Practice Alone
II-A	Auditory Instruction/ Practice Increases Print/ Braille Reading Speed
II-B	Auditory Instruction + Practice > Practice Alone (Print/ Braille Reading)
II-C	Auditory Practice + Print/ Braille Reading > Auditory Instruction and/ or Practice Alone (Print/ Braille Reading).

TABLE 27.6

MEANS, GILMORE ORAL READING TEST, ACCURACY SCORE

Group	Reading Mode	No.	Means Pretest	Posttest
1	Braille	6	85.5	87.8
	Print	5	88.0	95.0
	Combined	11	86.6	88.9
2	Braille	8	88.5	91.0
	Print	4	90.5	96.0
	Combined	12	89.2	90.0
3	Braille	8	84.5	88.8
	Print	3	85.3	87.3
	Combined	11	84.7	88.4
4	Braille	9	83.2	87.4
	Print	2	91.5	99.0
	Combined	11	85.3	89.5

The last three Listening Comprehension tests were taken from material that had been recorded and compressed from the reading text used for the practice sessions. All of these passages had been read by a member of the faculty of the Division of Radio, Television, and Motion Pictures of the University of North Carolina at Chapel Hill, recorded in their studios and compressed by the Center for Rate-Controlled Recordings at the University of Louisville.

The controlled group took two of these tests: one compressed at 20% and one compressed at 50%. The results of these two tests will be discussed, therefore. Analysis of variance (see Appendix Table 27.3) indicated significant effects for groups, reading mode, and interaction between groups and reading mode, as well as interaction among Ss by group by reading mode by tests.

Group 3 was again superior to each of the other groups. Group 2, however, did not perform significantly better than Group 1 as had been anticipated. Both Group 1 and Group 2 did attain scores higher than the controlled group, and these differences were significant. The means can be seen in Table 27.7.

TABLE 27. 7

MEANS, LISTENING COMPREHENSION TESTS 4 AND 6

Test &	Group 1		Group 2		Group 3		Group 4	
Reading Mode	M	S. D.	M	S. D.	M	S. D.	M	S. D.
Braille readers								
No.	6		8		8		9	
Test 4 (20%)	57. 2	3. 4	50. 1	7. 5	56. 0	7. 6	46. 3	9. 9
Test 6 (50%)	49. 5	7. 6	47. 2	9. 3	55. 8	11. 0	48. 1	8. 2
Combined	53. 3	7. 0	48. 7	8. 5	55. 9	9. 4	47. 2	9. 1
Print readers								
No.	5		4		3		2	
Test 4 (20%)	37. 6	7. 9	53. 0	7. 1	51. 7	8. 2	41. 5	12. 5
Test 6 (50%)	49. 2	9. 5	53. 0	8. 7	44. 3	16. 6	46. 0	3. 0
Combined	43. 4	10. 5	53. 0	8. 0	48. 0	7. 8	43. 8	9. 4
Combined								
No.	11		12		11		11	
Test 4 (20%)	48. 3	11. 4	51. 1	7. 6	54. 9	8. 0	45. 4	15. 9
Test 6 (50%)	49. 4	8. 1	59. 2	9. 5	53. 6	11. 1	47. 7	7. 7
Combined	48. 8	10. 8	50. 1	8. 6	53. 7	9. 7	46. 6	9. 5

The standard error of L is estimated as follows: $\sqrt{\lambda s^2 (X^2 / n_L)}$ with degrees of freedom obtained from the "s."

Except in Group 1, braille readers performed better than print readers. The difference between braille and print readers was significant on the compressed edition but not on the uncompressed edition ($\underline{t} = 3.631$, $p < .01$, 37 df). The means can be seen in Table 27. 8.

In studying the results of the short comprehension test administered after each practice session (see Appendix Table 27. 4), it was noted that after each increment in speed at 10% and 20% compression, there was a slight comprehension decrease on first exposure, followed by a subsequent recovery. At 30% and 50%, there was some indication that more exposure was needed if this recovery was to occur. On only two of these sessions were significant group differences found.

On the Gilmore Oral Reading Test, both on measures of accuracy and comprehension, no differences were seen that could be attributed to the training programs. The ANOVA of the accuracy measure can be seen in Appendix Table 27. 5. The ANOVA and means for comprehension are shown in Appendix Tables 27.6 and 27.7. Other educational experiences apparently effected the gains that were noted since they are seen in the controlled group as well as the experimental groups.

TABLE 27.8

MEAN SCORES STEP LISTENING TEST, FORM 2A AND 2B

Test &	Group 1		Group 2		Group 3		Group 4	
Reading Mode	M	S.D.	M	S.D.	M	S.D.	M	S.D.
Braille readers								
No.	6		8		8		9	
2A	277.2	9.5	289.9	16.7	295.9	13.8	289.1	11.1
2B	282.7	5.1	285.9	15.2	296.8	16.3	281.0	13.9
Combined	280.7	7.9	287.9	16.1	296.3	15.1	285.0	13.4
Print readers								
No.	5		4		3		2	
2A	279.4	5.0	282.2	6.8	289.3	17.9	287.0	13.0
2B	283.4	10.3	279.8	9.1	285.7	7.6	283.5	1.5
Combined	281.6	8.4	281.0	8.2	287.5	13.9	285.2	9.4
Combined								
No.	11		12		11		11	
2A	279.1	11.5	287.3	15.9	294.1	7.8	288.7	15.5
2B	283.2	10.9	283.8	13.8	293.7	7.4	281.4	15.3
Combined	281.1	8.1	285.6	14.4	293.9	15.3	285.1	12.6

For oral reading rate, a drop was seen among braille readers. Among print readers in Group 2, there was a considerable increase in reading speed, but this was not found for the print readers in the other groups. These print readers in Group 2 attained significantly better scores.

Since there were also significant differences found for reading mode, differences between braille and print readers were examined. In Group 2 it can be noted that print readers are superior to braille readers, whereas the reverse is true for each of the other groups. Using two modalities for reading simultaneously (tactual or visual and listening) proved more effective among the print readers than braille readers. The same difference between braille and print readers in Group 2 is seen in some of the other tests.

On the STEP Listening Test, Level 2, two forms were used: one of these was compressed at 30% and the other was played at the regular recorded speed.

On these tests the ANOVA (see Table 27.9) indicated no significant effects for tests. Group and reading mode effects were found. Interactions between group and reading mode and among group by reading mode by tests were also seen.

TABLE 27.9

ANOVA, STEP LISTENING TEST, FORM 2A AND 2B

Source of Variance	df	SS	MS	F
Total	89	16,955.172		
Group	3	1,919.189	639.729	37.178**
Reading Mode	1	413.388	413.388	24.025**
Group x Reading Mode	3	175.350	58.450	3.397*
Subjects (Group x Reading Mode)	37	599.663	17.207	
Tests	1	72.900	72.900	
Group x Test	3	376.378	125.459	
Reading Mode x Test	1	35.037	35.037	
Group x Reading Mode x Test	3	11,047.045	3,682.347	58.823**
Subjects (Group x Reading Mode x Test)	37	2,316.222	262.600	

*$p < .05$

**$p < .01$

On these tests Ss in Group 3 performed better than those in any of the other groups, thus support for primary hypothesis 1-A, instruction + practice increases auditory reading efficiency, was found.

Comparison of mean differences made throughout are based on the procedure outlined by Snedecor and Cochran (1967, pp. 268-279). According to this formula, t = the differences between the means under consideration divided by the standard error of 1, where L is any linear combination so that

$$L = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + \dots \lambda_k x_k$$

On the SORT, which consists of lists of words at graded difficulty levels, the tests differences between pre- and postperformance were not significant. The means did show improvement but these gains were found in the controlled group as well as the experimental group and, again, it was assumed that even had the differences proved significant it could not have been attributed to the experimental training program.

On the other measure of tactual or visual reading speed, the establishment of silent reading rate effects, the reading mode differences were significant and interaction among groups by reading by tests was seen (see Table 27.10).

TABLE 27.10

ANOVA SILENT READING RATE, PRE- AND POSTTRIALS

Source of Variance	df	SS	MS	F
Group	3	7,227.567	2,409.186	
Reading Mode	1	30,784.189	10,784.189	6.870*
Group x Reading Mode	3	1,027.896	342.632	
Subjects (Group x Reading Mode)	37	165,798.160	4,481.031	
Test	1	199.362	199.362	
Group x Test	3	618.774	206.258	
Reading Mode x Test	1	21.228	21.228	
Group x Reading Mode x Test	3	24,825.902	8,275.292	27.095**
Subjects (Group x Reading Mode x Test)	37	11,300.356	305.415	

*$p < .05$
**$p < .01$

When the differences between braille and print readers were examined in the three experimental groups, these were significant. Print readers in Group 2 did show significant gains in silent reading rate just as they did in oral reading rate. This suggests that benefits from the listening-reading training resulted for print readers using two modalities. The gains for these readers averaged 32.2 words per minute ($\underline{t} = 2.164$, $p < .01$, 37 df).

Except for the print readers in Group 2, there is little evidence of reading gains in visual or tactual reading in the other groups. When two modalities are used simultaneously, print readers will receive benefits in certain respects apparently, but no other group of Ss seem to experience any subsidiary gains.

On the Brown-Carlsen Test of Listening Comprehension, no significant main effects were revealed by the analysis of variance by any of the subtests. Both the subtests on recognizing transition and recognizing words in context would seem to be measures of visual or tactual reading skill just as much as listening skill. The training program appeared to contribute little to these tasks.

The Lecture Comprehension subtests measure a listening situation where the material heard was prepared for oral delivery rather than a listening-reading task. It may have been inappropriate in this particular study but it was employed since it is one of the few tests available at a listening ability. How efficient an instrument this was in prediciting performance on the other comprehension measures has not been examined. It was not, apparently, efficient in measuring the effects of listening-reading training.

Summary

Out of the 45 tests of the primary hypotheses, 27--or 60%--give positive support to the first three hypotheses. There is evidence that listening technique training followed by practice using compressed speech as an aid is an efficient method of effecting gains in listening-reading. These gains are apparent both with uncompressed and moderately compressed materials.

Six of the 45 tests supported the hypothesis that use of two modalities simultaneously in reading would cause improvement in listening-reading. Only three tests gave evidence of improvement that occurred from listening practice only.

Only for print readers in Group 2, who listened while they read the passages, was there any evidence that the training had positive effects on visual or tactual reading. Since it was noted that as speeds became greater, braille readers found it increasingly difficult to keep up with what they heard, and that these readers abandoned the attempt to do so after a time, it would be reasonable to assume that braille readers would need more exposure at expanded speeds nearer their reading rate, with gradual increases in these, if they are to experience gains in braille reading speed. The results would indicate that it would be worthwhile exploring this. As far as improving listening-reading, using two modalities at the same time was not as effective as giving listening training.

Practice with listening-reading, using compressed speech at graduated increments and followed by short comprehension tests, was the most inefficient method. Visually impaired students, by the high school age level, have had considerable experience with reading by listening at normal recorded speed. Practice has been available to them over a period of time. More practice, even when using compressed speech rates, does not seem to be as efficient as training them in better techniques of listening.

One interesting finding that had not previously been anticipated was that a difference existed between braille and print readers in performance on almost every test administered at the conclusion of the training program. Braille readers tended to perform better on the tests than did print readers except insofar as reading rate was concerned. When it is taken into account that these groups were not significantly different except in reading rate, this finding becomes quite interesting.

Whether this difference in performance, which shows braille readers to be able to utilize training for improvement of listening more effectively, stems from greater motivation or from greater development of cognitive capacities in utilizing materials learned from auditory reading, or a combination of these factors, is not clear. The print readers were all severely visually impaired and will necessarily depend on auditory input in many situations where individuals with normal vision use visually acquired information. Their need for learning to utilize auditory skills is great, but recognition of this may not be present and there may be a rejection of this as signifying "blindness" among some individuals. The student without useful vision for reading, no matter how much he may react to being visually impaired, is forced to accept the fact that he must depend on listening for many reading experiences. In working with partially impaired individuals, who retain useful vision, functioning in a public school setting who have had more experience with the need for listening-reading, there might conceivably be a difference in recognition of need for auditory reading improvement.

In a residential school the student is more likely to find texts in braille than he would be in a program in a public school. The student who reads visually is more likely to have available to him the large print edition of the text than would be true in public school settings many times where the particularly needed text is not available.

On the other hand, the student who is so severely visually impaired that vision cannot be a channel for acquiring more than minimum information about light or shape has also necessarily been put in a position where he must use the auditory channel more extensively at the conscious level. He has thus exercised and developed these capacities in his previous experiences. For this reason he may be in a better position to utilize training and practice designed to increase these abilities both from the standpoint of educational and psychological readiness.

It might be noted that participants enjoyed the book that was read. Some selections were naturally received more favorably than others, but interest remained high throughout the sessions. After each increase in speed, students would complain mildly that they could not understand it, but by the second exposure they usually indicated that it was a desirable rate up through 30% compression. Fifty percent compression continued to be difficult to understand. A few students in considering their reaction to it suggested that after a little further use it was clear what was being said, but they felt so rushed they had difficulty reacting to it and therefore could not retain it. A few students, who normally read braille or print more rapidly than others, did attain scores on the comprehension test and questions after practice sessions that suggest they were able to understand and retain 75% or more of what was heard. Since, of course, more material was covered, there was more to be mastered at these faster speeds. The very increase in material to be learned put the student with less academic aptitude at a disadvantage. The students indicated a choice of speeds for future reading if they were able to specify compression in the future. This seemed to be almost equally divided between 20% and 30% as a desirable choice under these conditions.

Utilization of auditory reading can be improved by instruction. This type of instruction should probably be incorporated systematically into teaching programs for the sensorially and perceptually impaired among the visually handicapped, but it should not be necessarily limited to usage with this group. There are implications for general education programs and programs for the educationally disadvantaged in using compressed speech. These have been largely untapped in any systematic manner although the volume of available materials has increased and is increasing.

APPENDIX

APPENDIX TABLE 27.1

ANOVA, LISTENING COMPREHENSION TESTS 1, 2, AND 3

Source of Variance	df	SS	MS	F
Total	102	9,759.180		
Group	2	1,613.984	806.992	18.002**
Reading Mode	1	11.672	11.672	
Group x Reading Mode	2	251.330	125.665	
Subjects (Group x Reading Mode)	28	1,288.197	44.828	
Test	2	8.015	4.008	
Group x Test	4	289.135	72.284	
Reading Mode x Test	2	615.496	307.748	4.390**
Group x Reading Mode x Test	4	1,788.964	447.241	6.380**
Subjects (Group x Reading Mode x Test)	56	3,925.386	70.096	

**p < .01

APPENDIX TABLE 27.2

MEAN SCORES, LISTENING COMPREHENSION TESTS 1, 2, AND 3

Test	Group 1 Mean	Group 1 S.D.	Group 2 Mean	Group 2 S.D.	Group 3 Mean	Group 3 S.D.
Braille readers						
No.	6		8		8	
1	44.2	11.3	48.2	7.1	53.6	5.8
2	42.7	8.6	43.1	11.5	58.0	5.6
3	51.7	6.9	49.5	5.4	57.8	8.1
Combined	46.2	6.5	46.9	8.8	56.4	6.9
Print readers						
No.	5		4		3	
1	46.8	12.5	56.5	4.4	49.3	11.5
2	51.4	5.0	52.5	7.9	54.3	5.9
3	38.8	8.5	46.0	13.1	55.0	3.0
Combined	45.7	15.5	51.7	9.3	52.9	6.8
Combined						
No.	11		12		11	
1	45.4	11.9	51.0	7.5	52.4	7.2
2	46.6	8.4	46.3	11.1	57.0	5.6
3	45.8	10.1	48.3	8.4	57.0	7.5
Combined	45.9	10.2	48.5	14.4	55.5	11.8

APPENDIX TABLE 27.3

ANOVA, LISTENING COMPREHENSION TESTS 4 AND 6

Source of Variance	df	SS	MS	F
Total	89	8,609.379		
Group	3	590.044	196.681	11.007**
Reading Mode	1	190.250	190.250	10.647**
Group x Reading Mode	3	757.017	252.339	14.122**
Subjects (Group x Reading Mode)	37	661.151	17.869	
Test	1	1.111	1.111	
Group x Test	3	81.815	27.272	
Reading Mode x Test	1	214.241	214.241	
Group x Reading Mode x Test	3	3,362.901	1,120.967	15.078**
Subjects (Group x Reading Mode x Test)	37	2,750.849	74.347	

**p < .01

APPENDIX TABLE 27.4

MEAN PERCENTAGE OF CORRECT RESPONSES, PRACTICE SESSION QUESTIONS FOR EXPERIMENTAL GROUPS

Session	Percent Compression	Mean Percentage Correct Responses Group 1	Group 2	Group 3
Test				
1	0	77.8	77.8	74.4
2	10	45.4	54.2	63.6
3	10	77.3	75.0	77.3
Test				
5	20	63.6	63.9	81.8
6	20	61.2	67.4	70.9
Test				
8	20	61.4	56.1	60.0
9	20	47.5	62.9	70.9
Test				
11	30	39.8	54.5	45.4
12	30	55.0	57.6	63.4
13	30	53.0	53.5	50.0
Test				
14	50	42.5	50.0	41.7
15	50	44.2	37.1	45.4

APPENDIX TABLE 27.5

ANOVA, PRE- AND POSTTEST RESULTS, ACCURACY SCORE, GILMORE ORAL READING TEST

Source of Variance	df	SS	MS	F
Group	3	253.640	84.550	7.155**
Reading Mode	1	351.101	351.101	29.712**
Group x Reading Mode	3	240.400	80.133	6.781**
Subjects (Group x Reading Mode)	37	437.234	11.817	
Test	1	352.044	352.044	2.235
Test x Group	3	481.346	160.448	
Test x Reading Mode	1	829.533	829.533	5.266**
Test x Group x Reading Mode	3	83.149	27.716	
Subjects (Test x Group x Reading Mode)	37	5,828.738	157.533	

**p < .01

APPENDIX TABLE 27.6

MEANS, GILMORE ORAL READING TEST, ACCURACY SCORE

Group	Reading Mode	No.	Means Pretest	Means Posttest
Group 1	Braille	6	85.5	87.8
	Print	5	88.0	95.0
	Combined	11	86.6	88.9
Group 2	Braille	8	88.5	91.0
	Print	4	90.5	96.0
	Combined	12	89.2	90.0
Group 3	Braille	8	84.5	88.8
	Print	3	85.3	87.3
	Combined	11	84.7	88.4
Group 4	Braille	9	83.2	87.4
	Print	2	91.5	99.0
	Combined	11	85.3	89.5

APPENDIX TABLE 27.7

MEANS, GILMORE ORAL READING TEST, COMPREHENSION SCORE

Group	Reading Mode	No.	Means Pretest	Means Posttest
Group 1	Braille	6	44.8	44.3
	Print	5	46.4	46.6
	Combined	11	45.6	45.4
Group 2	Braille	8	45.6	46.4
	Print	4	45.2	46.5
	Combined	12	45.5	46.4
Group 3	Braille	8	45.6	47.5
	Print	3	43.3	45.3
	Combined	11	45.0	46.9
Group 4	Braille	9	45.4	45.0
	Print	2	45.5	45.5
	Combined	11	45.4	45.1

APPENDIX TABLE 27.8

ANOVA, GILMORE ORAL READING TEST, COMPREHENSION SCORE

Source of Variance	df	SS	MS	F
Group	3	3.623	1.208	
Reading Mode	1	.269	.269	
Group x Reading Mode	3	19.501	6.500	
Subjects (Group x Reading Mode)	37	509.562	13.772	
Test	1	6.013	6.013	
Group x Test	3	5.319	1.440	
Reading Mode x Test	1	1.085	1.085	
Group x Reading Mode x Test	3	.572	.191	
Subjects (Group x Reading Mode x Test)	37	138.712	3.749	

APPENDIX TABLE 27.9

ANOVA, GILMORE ORAL READING TEST, ORAL READING RATE

Source of Variance	df	SS	MS	F
Group	3	1,512.041	504.013	
Reading Mode	1	3,873.515	3,873.515	4.151*
Group x Reading Mode	3	470.380	156.793	
Subjects (Group x Reading Mode)	37	34,525.604	933.124	
Test	1	421.764	421.764	2.78
Group x Test	3	398.311	132.770	
Reading Mode x Test	1	162.587	162.587	
Group x Reading Mode x Test	3	446.351	148.784	
Subjects (Group x Reading Mode x Test)	37	5,597.638	151.288	

*p < .05

APPENDIX TABLE 27.10

MEAN WPM, GILMORE ORAL READING TEST

Group	Reading Mode	No.	Means Pretest	Means Posttest	Posttest Mean - Pretest Mean
Group 1	Braille	6	80.2	61.9	-18.2
	Print	5	99.6	92.9	- 6.7
Group 2	Braille	8	89.5	78.1	-11.4
	Print	4	101.1	120.1	+19.0
Group 3	Braille	8	71.8	66.2	- 5.6
	Print	3	97.5	86.9	-10.6
Group 4	Braille	9	82.9	73.8	- 9.1
	Print	2	93.9	79.8	-14.1
Total	Braille	31	81.2	70.6	-10.6
	Print	14	98.8	97.5	- 1.3

CHAPTER XXVIII

COMPREHENSION FOR IMMEDIATE RECALL OF TIME-COMPRESSED SPEECH AS A FUNCTION OF SEX AND LEVEL OF ACTIVATION OF THE LISTENER

Sally R. McCracken*

Clive Lewis once wrote: " . . . the future is something which everyone reaches at the rate of sixty minutes an hour, whatever he does, whoever he is." The purpose of this study was to determine the sex-related similarities or differences of comprehension scores and Galvanic Skin Response (GSR) measurements for a normal rate of recorded speech (160 words per minute [wpm]) and a time-compressed recorded speech (320 wpm).

Definition of Terms

Recorded normal rate speech was defined for this study as that reading rate deemed necessary, by a professional radio announcer,** to achieve the greatest degree of clarity and understanding for the listener. With these instructions, the reader averaged approximately 160 wpm. Time-compressed speech was defined as a method of shortening the playback time of recorded materials without change of pitch or loss of the original information (Foulke, 1967b). The word rate was approximately 320 wpm. Rapid recall in terms of listening comprehension was defined as the ability to recall information presented in selections immediately following presentation as measured by means of a multiple-choice test. Level of activation was defined as " . . . the extent of release of potential energy stored in the tissues of the organism, as this is shown in activity or response [Duffy, 1962, p. 64]." (GSR)

---

*Dr. Sally R. McCracken is an Assistant Professor of Speech and Dramatic Arts, Eastern Michigan University, Ypsilanti, Michigan 48197.

**J. Daniel Logan, announcer for WCAR and WDET, Detroit, Michigan.

Procedure

Ninety sighted males and 90 sighted females, between the ages of 19 and 22, were matched according to grade point ratio, screened by an auditory acuity test, equipped with headsets, and exposed to eight different readings selected from the Diagnostic Reading Tests, Section II, Comprehension Silent and Auditory. Four of the readings were presented at a normal rate and four at a compressed rate. The experimental design was based on the Wason Model (Wason, 1962) which enabled each S to act as his own control. The original Wason design appeared as Ab, bA, aB, and Ba indicating that each "A" or full-length selection has an "a" or shorter version. The design was interpreted for this study to mean that for the first four selections on each tape every normal version or "A" had a compressed version or "a." The last four selections (five through eight) were designated as "B" for the normal rate and "b" for the compressed rate. The design for the main study appeared visually as:

	Main Order			Order Control	
Story	Tape 1	Tape 2	Story	Tape 3	Tape 4
1	A	a	5	b	B
2	A	a	6	b	B
3	A	a	7	b	B
4	A	a	8	b	B
5	b	B	1	A	a
6	b	B	2	A	a
7	b	B	3	A	a
8	b	B	4	A	a

During the testing period, each S was measured by means of the GSR. Care was taken to avoid excessive movement during the actual selection listening periods. Following each of the eight readings, the S was tested for comprehension of the recorded material. The session lasted approximately 30 minutes. Before leaving the testing area, each S completed a written reaction sheet concerning time-compressed speech.

Testing Objectives

The testing procedure was structured to answer the following questions.

1. When listening to a normal rate of recorded speech is there a relationship between the listening comprehension scores and the level of activation or physiological arousal of the listener? (GSR)

2. When listening to a normal rate of recorded speech is there a relationship between the listening comprehension scores and the sex of the listener?
3. When listening to recorded time-compressed speech is there a relationship between the listening comprehension scores and the level of activation or physiological arousal of the listener? (GSR)
4. When listening to recorded time-compressed speech is there a relationship between the listening comprehension scores and the sex of the listener?
5. Are there differences and/or similarities between the experiences of listening to a normal rate of recorded speech or recorded time-compressed speech in terms of listening comprehension scores, level of activation or physiological arousal (GSR), and the sex of the listener?

Data Preparation

The comprehension scores were compiled and a computerized formula for conversion of the GSR scores was prepared to provide a basis for correlation and analysis of variance (Lacy & Siegel, 1949). The conversion formula (Lacy & Siegel, 1949) accounted for basal and amplitude variance since some Ss entered the testing session at a higher level of activation than other Ss. Resistance ohm measurements were converted to conductance micromhos in the formula $C = (1/R \times 10^6) - (1/R \times 10^6)$. These data were programmed and submitted for computer analysis to determine the interrelationships among the measurements.

Results

1. Listening comprehension scores and GSR scores appeared to represent two distinctly different phenomena and did not correlate for either the compressed rate or the normal rate selections (Table 28.1).
2. There were no sex-related differences in comprehension scores for the normal rate of speech.
3. Listening comprehension scores and GSR scores did not correlate when the stimulus factor was time-compressed speech.
4. No sex-related differences were found concerning the listening comprehension scores when the stimulus factor was time-compressed speech (Table 28.2).
5. When the level of compression was held to 50% the listening comprehension scores were identical to the normal rate scores for both sexes. An interesting effect did occur when males tended to register higher GSR scores when exposed to time-compressed speech than did the females. There was also a tendency for the males to react at a higher level to the normal rate presentations. Figure 28.1 illustrates male and female

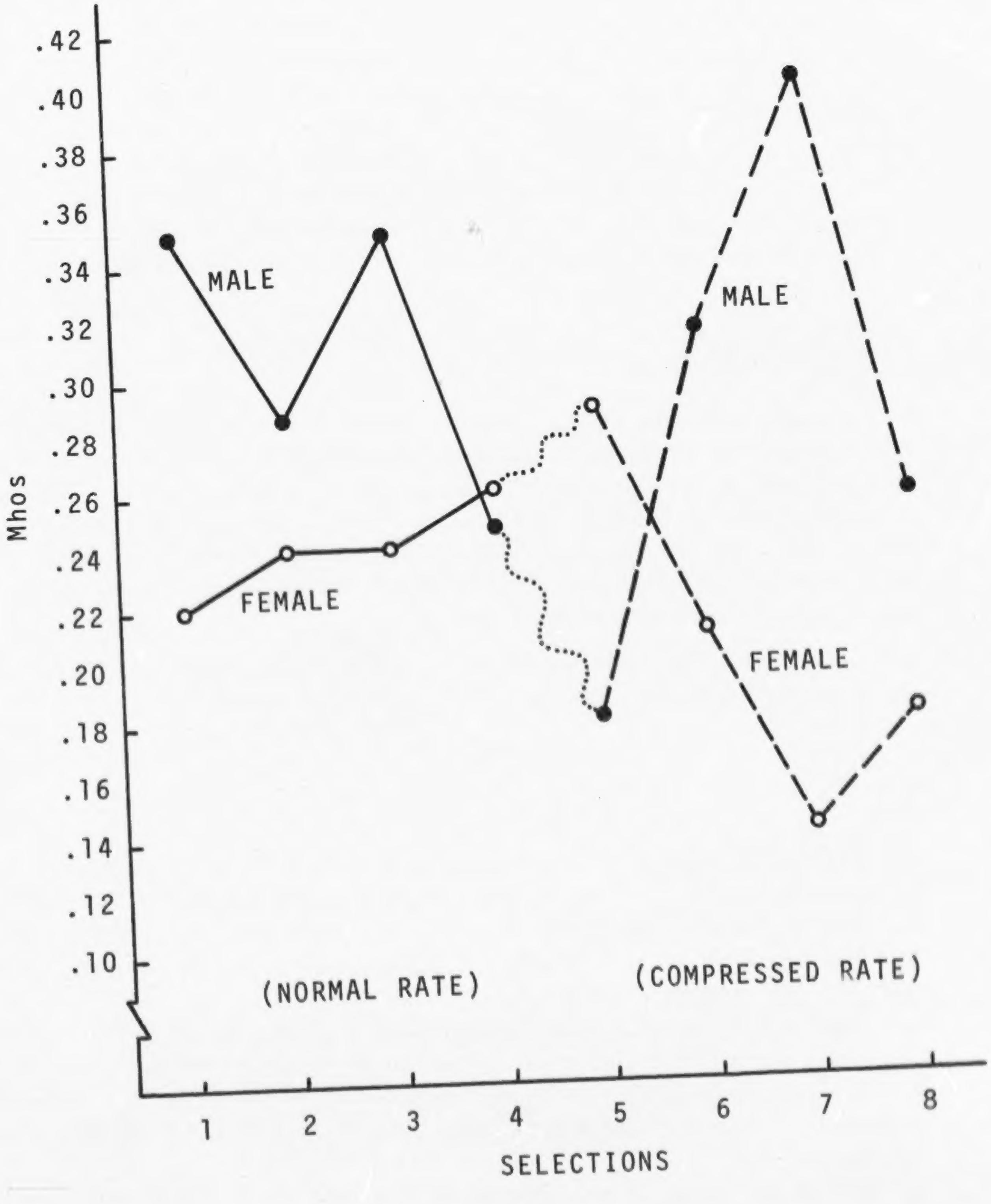


Figure 28.1. Males and females, normal rate followed by compressed rate.

reactive tendencies when exposed to four normal rate selections and four compressed rate selections.

TABLE 28.1

CORRELATIONS FOR MALES AND FEMALES FOR LISTENING COMPREHENSION, RESISTANCE GSR (ROHM), CONDUCTANCE GSR (MHO), REACTION TO NORMAL RATE (REAN), AND REACTION TO COMPRESSED RATE (REAC)

		N	ROHM	MHO	REAN	REAC
Comph.	M&F	180	-.0219	.0103	.0758	.1629
	M	90	.0098	-.0116	.0172	.1895
	F	90	-.0438	.0300	.1444	.1405

TABLE 28.2

LISTENING COMPREHENSION SCORES FOR MALES AND FEMALES, NORMAL RATE AND COMPRESSED RATE

	N	$\overline{X}$	S.D.	$\overline{X}$df	t	SIG.
Males	90					
Normal	90	19.0888	7.9918	1.0222	.5000	N.S.
Compressed	90	18.0666	8.6075			
Females	90					
Normal	90	18.6888	9.8680	.4888	.3636	N.S.
Compressed	90	18.2000	8.1470			

Figure 28.2 illustrates males and females reactive tendencies when exposed to four compressed rate selections followed by four normal rate selections.

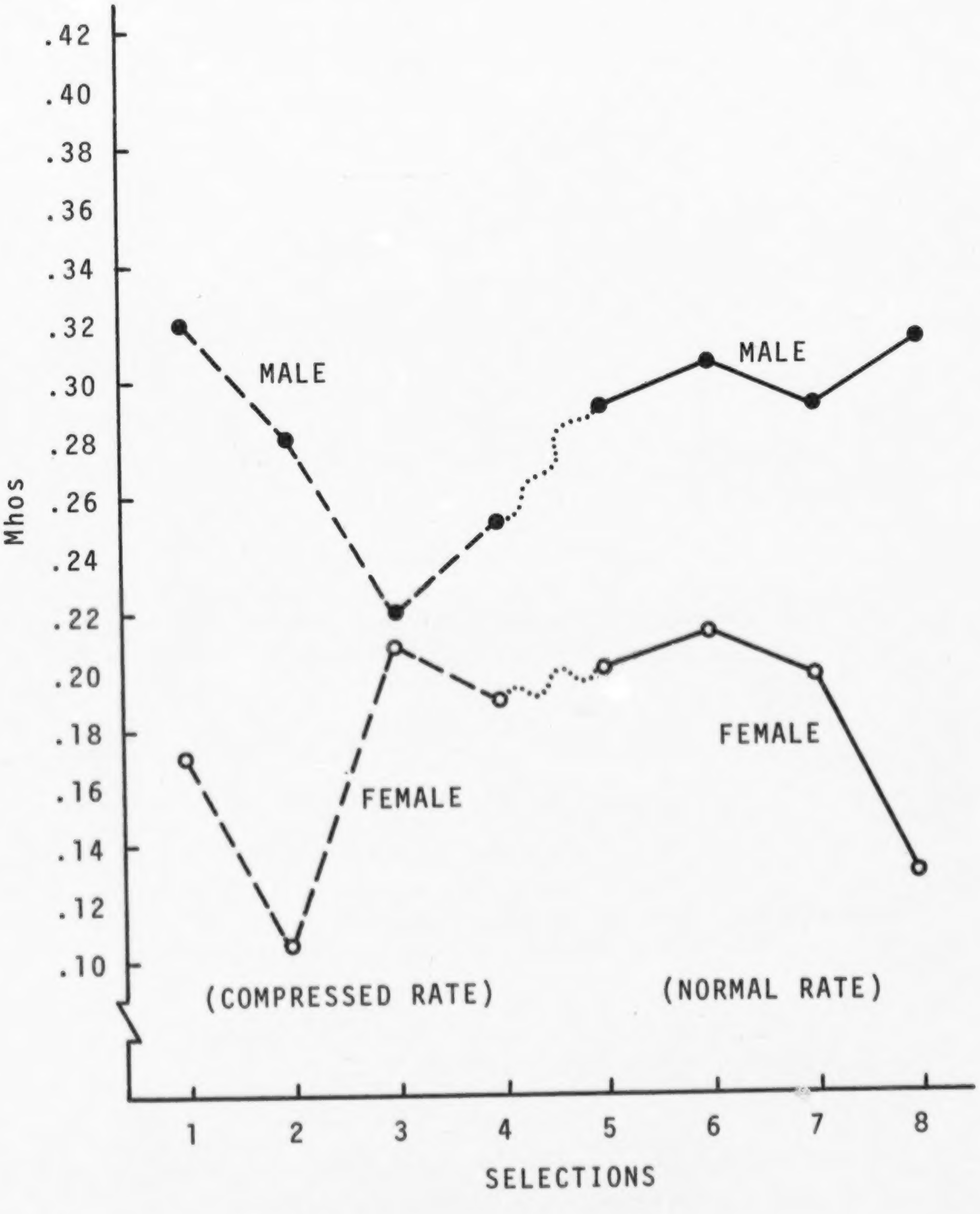


Figure 28.2. Males and females, compressed rate followed by normal rate.

Discussion of the Results

The finding that the listening comprehension scores were equal for the normal and the compressed rates supported the results of Fairbanks, Guttman, and Miron (1957c); Foulke, Amster, Nolan and Bixler (1962); and Orr, Friedman, and Williams (1965). These studies predicted a slight loss of listening comprehension for the 50% level but the researchers tended to suggest that this was justifiable when the consideration of saved time was realized. The present study did not find any loss in comprehension due to the 50% level of compression.

Sex-related differences for comprehension ability did not appear to be a major concern. Males and females comprehended the material equally well and the stimulus recordings were not intentionally sex biased. The groups were matched with great care which could account for a lack of sex difference in comprehension scores.

According to the present study, the males tended to effect higher levels of activation when listening to time-compressed speech at a 50% level. The literature provided no information as to the possible physiological reaction levels for males or females when listening to time-compressed speech. It is interesting to note that the written reactions of the Ss indicated a "stress" experience when listening to time-compressed materials, but this was not substantiated by the GSR measurement. The females, more often than the males, tended to indicate they experienced a "stress" increase but did not physiologically register this feeling according to the GSR scores. The males expressed less "stress" in the written reactions yet effected higher GSR scores. It appears that the reactions of the females tended to be the result of a mental set rather than an actual physiological strain. They could have reasoned that ". . . because the rate is faster it is more difficult . . . "* and therefore experienced a psychological increase in tension. The cause for males effecting higher GSR scores is left to speculation since the conversion formula accounts for differences in basal activation levels.

At the present time it is the opinion of this researcher that the felt stress experienced by the Ss was probably due to the mental factors and not to the actual experience of listening to time-compressed speech at a 50% level.

---

*Female, I. D. No. 323708. Written comment following the testing session.

CHAPTER XXIX

THE COMPREHENSION OF RATE-CONTROLLED SPEECH BY SECOND-GRADE CHILDREN WITH FUNCTIONAL MISARTICULATIONS

R. Vernon Stroud*

The purpose of this study was to ascertain if rate of speaking affected the comprehension of speech by second-grade children with functional misarticulations. It was also the purpose of the study to ascertain the effects of sex, severity of misarticulations, therapy, socioeconomic status, size of family, and race on the ability to comprehend speech altered by various degrees of speaking rates.

However, in the paper presented here, concern was placed upon those children with functional misarticulations.

Ninety-eight second-grade children from the Dayton, Ohio, Public School System served as Ss. Fifty-two of these children were normal speakers and 46 were diagnosed by the speech clinician as having functional misarticulations. All of the Ss who possessed normal hearing were average or above in intelligence.

Each S was individually exposed to stimuli which consisted of rate-controlled speech which was in the form of imperative and interrogative sentences. One hundred and seventy sentences were used as stimuli. The various rates of speech used in the study were accomplished through the use of a Tempo Regulator which compressed (speeded) and expanded (slowed) the speech electronically.

The sentences used as stimuli were divided into units of 10, and each unit was processed by the Tempo Regulator. The first unit which was the reference point was presented at 225 words per minute (wpm) and 280 syllables per minute (spm). The second unit received a 10% increment and was presented at 248 wpm and 308 spm. This procedure continued until 383 wpm and 476 spm were reached. The procedure for the expanded (slow) stimuli was similar to the one used in the compressed condition except the stimuli received a 10% decrement in relation to the reference point which was 225 wpm and 280 spm.

---

*Dr. R. Vernon Stroud is affiliated with the Barney Children's Medical Center, Dayton, Ohio 45404.

In that the results of the expanded study were insignificant, they will not be reported here.

Casual observation has led this investigator to feel that some children may miss part of the oral message or phonemic differences of spoken language as a function of the rate of talking. Unfortunately, search for literature in which this observation is quantified has been unrevealing. Some investigators have substantiated that children with functional misarticulations have difficulty hearing phonetic elements in words and the same has been reported with reference to the child's ability to hear phonemic differences and similarities in words. However, to this investigator's knowledge, no study in which the objective was to observe the perceptual abilities of children with misarticulations in relationship to rapid speech has been reported.

A relevant and interesting point of view is the one presented by Liberman (1961). According to Liberman, the perception of speech does not depend solely on the acoustical characteristics of the stimulus; instead it is perceived in reference to articulation. He postulated that the articulator's movements and their sensory effects mediate between the acoustic stimulus and perception. In essence, the listener mimics the incoming message and responds to proprioceptive and tactile stimuli that are produced by his own articulatory movements. This permits one to infer that if a listener is unable to perform these processes as a function of the rate of utterance, he will have significant difficulty comprehending the message.

With this in mind, this investigator postulates that children with functional misarticulations would become so bogged down in attempting to associate sound with place of articulation that they would fail to comprehend the spoken message. In order to do this, two groups of children had to be studied. One group consisted of 52 normal speaking second-grade children; the other consisted of 46 second-grade children who were reported as having functional misarticulations. Both groups were described as having normal hearing and being of normal intelligence on the basis of stanine scores obtained on The California Mental Maturity Test.

The next task was to subject these children to speech which had been compressed. Speech was compressed at the following rates:

WPM: 225, 248, 270, 293, 315, 338, 360, 383

and

SPM: 280, 308, 336, 364, 392, 420, 448, 476

This was presented to the Ss via the Wollensak model 1500 tape recorder. The stimuli consisted of imperative and interrogative sentences which had been compressed on a Tempo Regulator at The Center for Rate-Controlled Recordings, University of Louisville, Louisville, Kentucky. Ten sentences were presented at each rate and the child was required to respond. For example, the voice on the tape would ask, "Is your teacher tall or short?" and the child would respond, "Tall." Another example, "Is it cold outside when Christmas comes?" and the child would respond, "Yes." At this point, the examiner would reply, "Yes, what?" and the child would say, "It is cold when Christmas comes." Five seconds were allowed as response time. If no response occurred after this period of time, the examiner would say, "Listen." and go to the next sentence. The errors made by both groups were recorded and treated statistically.

The results imply that children with functional misarticulations have greater difficulty in comprehending rapid speech than children with normal speech. The responses of the two groups were equal through 270 wpm or 336 spm. At 293 wpm or 364 spm, the defective speakers made their first significant errors. From this point on through 383 wpm or 476 spm, the defective speakers made significantly more errors than the normal speakers.

The findings of this study have many implications relative to the processing of information by children with functional misarticulations. Somewhere in these data or in future studies, one may find a clue to causality of articulation disorders.

CHAPTER XXX

RATE-CONTROLLED SPEECH AND SECOND LANGUAGE LEARNING

Herbert L. Friedman and Raymond L. Johnson*

I'd like to begin by indicating the ways in which rate-controlled speech is a technique which is relevant to second language learning, i.e., the acquisition of proficiency in a foreign tongue. Its relevance is primarily, I believe, to a phenomenon central to language learning which seems to me to be somewhat neglected at this conference, the phenomenon of listening behavior. I know we're deeply involved in discussions of speech rate and other major parameters of the speech stimulus, but what are we measuring? The nature of the work the listener *may* perform on a stream of speech includes such diverse activities as identification, discrimination, matching, storing for short, long, or intermediate durations, reiterating, paraphrasing, translating, comprehending the core meaning, associating to, anticipating, and so on. I don't say that all of these activities necessarily occur, only that they may occur, and that they may occur *during* the flow of speech. We know that, because we can evoke the nature of the information resulting from those activities in the listener. *What* we measure, and equally important what we ask the listener to do, are inescapable determinants of the conclusions we draw about effective or efficient listening.

Well, we don't know much about the nature and relationships of those delicately timed activities which constitute listening behavior. We have all correctly guessed, however, that when the time available to do work concurrent with speech is reduced, successful listening will not occur. By systematically reducing the time available to process speech without damaging intelligibility, i.e., by using time-compressed speech, we can learn what priorities the listener (consciously or otherwise) assigns to what he hears. By manipulating the restoration of time to coincide or jar against linguistic rules, we can further determine what the missing time is used for. Those are two now standard ways to manipulate the temporal characteristics of speech--but we ought also to look then at the latency with which a listener responds where that is feasible since that is time too, and possibly at the duration of his response or the overall response time. In other words we ought to look at the temporal aspects of response alone, and, of course, one can add to that, the degree of accuracy of the response and the degree of confidence with which the listener makes it.

---

*Dr. Herbert L. Friedman is Director of the Communication Skills Research Program at the American Institutes for Research, 8555 Sixteenth Street, Silver Spring, Maryland 20910. Dr. Raymond L. Johnson is Research Scientist in the program.

What I've just said hasn't really been a digression from our efforts in the second language learning area. It is, rather, an outline of the approach and some of the components we have used in attacking the problems of second language learning, as well as the selective perception of native speech, which my colleague Raymond Johnson reported on yesterday.*

The student of a second language is in a situation which resembles in a major way that of the listener to compressed native speech. In a situation in which all the speech is intelligible and the vocabulary and syntax within his grasp, the speech may yet be too fast for him to process adequately. (It may be worth giving some attention to the possibility that this is also true of the disadvantaged native listener).

The work the student must perform on foreign speech is greater than the necessary minimum with his native tongue because, I believe, some degree of translation work is still necessary for him, and that means, of course, not only the possibility of some word-by-word translation, but a restructuring of the language. The student is not taught to translate but, at the early stages, I think it still occurs. In any case something does, since it takes him longer to perform.

We looked at three questions during the course of this project.
1. How does the selective perception of language differ when speech is compressed in time and when time is restored in preselected places?
2. How does the nature of the task assigned to the student affect the latency of his performance under compressed and noncompressed conditions?
3. Would exposure to gradually increasing rates of speech in the second language enable a student to listen better at normal rates?

The basic research studies devoted to the first two questions were performed on a population of Russian language students at three levels enrolled at Georgetown and American Universities in Washington, D.C.

---

*See Chapter XIII, "Temporal Spacing and the Comprehension of Time-Compressed Speech."

An earlier training study was performed on Russian and Vietnamese (Hanoi dialect) students participating in the 37-week aural comprehension courses at the Defense Language Institute in Monterey, California. The training study reported below was done on a group of Russian language students at Georgetown University.

Speech Manipulation

Two of the most recent stimulus manipulation studies performed, whose findings only I'll mention here because of the time constraint, were designed to help us identify priorities in the recall of sentences. In one case two 'string types' were employed after Miller and Isard (1963): grammatical meaningful and grammatical meaningless Russian sentences. In the second study all sentences were meaningful grammatical but some were segmented structurally by inserting temporal space between kernel and adjunct, some were nonstructurally segmented, and some were unsegmented. The sentences were so constructed by our Russian consultants that kernel and adjunct did not overlap. The overall findings from these studies (and some performed earlier) parallel our findings with native speech.

In the first study, not surprisingly, difficulty in recall is increased both by anomalousness and compression. The kernel portions of sentences are recalled with greater accuracy than the adjunct, indicating the importance of syntactic recognition. It may suggest that anomalousness makes syntactic recognition more difficult while compression deprives the listener of time to do it.

The greater the proficiency of the student (i.e., year of enrollment), the shorter is his latency of response. There is also a position effect, in that earlier portions of the sentence are better remembered than later ones. The insertion of temporal spaces in the second study produced highly significant results when the space was at structural locations. Nonstructural segmentation seemed to interfere with recall although performance wasn't significantly different from nonsegmented speech. Adjectives were lost more frequently than nouns. A common finding--but one reaffirmed here for all conditions which made a sentence more difficult.

In general then, second language listeners in our studies employed meaning preserving techniques and did so more frequently the more experience they had with the language.

Task Manipulation

In another study four tasks were assigned to students at two levels of proficiency in Russian. The tasks (performed on the last word of single conventional sentences) were: (1) simple repetition of Russian word, (2) translation of Russian word to English, (3) substitution of different but appropriate word in Russian, and (4) substitution of different but appropriate English word.

The complexity of the task, the level of proficiency of the student, and the rate of presentation were examined for the effects on response accuracy and response latency. It was hypothesized that for a simple task, or proficient student, (or both), the duration of the speech stimulus may be reduced by speech compression without affecting either latency or accuracy of response. Below a certain level of duration, however, the listener will have to restore some of the time removed to meet his minimal processing needs for accurate response. He can do that by increasing his latency without altering accuracy. A third stage (it was hypothesized) is reached when the initial sentence is presented so quickly that work which must be performed during the speech stimulus cannot be either accomplished or delayed until afterwards. At that point latency may be decreased or not affected, but accuracy will decrease.

The relationships hypothesized are highly complex and were certainly not all established in this study. However, the results do point to the likely validity of the overall paradigm. The findings indicated:
1. The tasks employed differ significantly from each other on either latency or accuracy measures.
2. Level of proficiency is most clearly indicated by latency rather than accuracy measures. An interesting exception to this occurs in a reversal of the effect, i.e., longer latency for more trained students, on a task which they have been trained not to do (English translation).
3. On the more complex tasks, increase in rate results in no major increase in latency, but it does lessen accuracy of performance.

Overall in both the task manipulation and stimulus manipulation studies, and indeed throughout all the basic research studies performed on this project, the findings repeatedly indicate that temporal variables are critical to the understanding of second language listening behavior both as they vary in the language and in the listener's deployment of work and priorities.

Training Technique

As well as its use as a basic research tool, rate-controlled speech may have considerable potential as a training device. In a short study recently completed, speech compression was joined with the added-parts technique of presenting material for efficient learning. In this technique a passage is incrementally expanded such that the first portion heard is heard on each subsequent presentation again, and with each repetition, a later portion of the passage is added. In this version Russian material presented initially at normal speed was presented at 1.3 x normal speed during its second playing, and 1.6 x normal speed at its third presentation; each new part being initially presented at normal speed.

Compared with a control condition in which speech compression was not used, there is no significant decrement in performance, suggesting that considerable time may be saved by using speech compression while achieving equivalent efficiency of training.

This paper, necessarily brief,* will I hope give you some sense of the diversity of rate-controlled speech as a basic and applied tool in second language learning as well as with native speech and, thereby, provoke further research in both areas.

ACKNOWLEDGMENT

This project sponsored by Advanced Research Projects Agency, Contract # DAHC 15 67 C 0262.

---

*For greater detail see Final Report, Contract # DAHC 15 67 C 0262, Advanced Research Projects Agency, December 31, 1969.

CHAPTER XXXI

COMPRESSED SPEECH IN MEDICAL EDUCATION

Gloria J. Boyle*

At the University of Missouri's School of Medicine, compressed speech has become useful in both the basic science and clinical areas of medical education (Figure 31.1) for several reasons:

1. Using compressed speech, research shows that learning is increased by forcing the student to listen conceptually and attentively.
2. Since the medical student is tightly scheduled, time is an especially valuable commodity. Through the use of compressed speech, material can be presented or reviewed at a faster rate, thus freeing the student for other learning activities.
3. In concentrated lecture presentations note-taking hinders the learning process. If the student is freed from this task with the reassurance of a compressed copy for later review purposes, more can be gleaned from the initial lecture.

In this context compressed speech is used for the first- and second-year students in their basic science education. Scheduled lectures are taped at regular speed and currently at 70% compression. Using the fast-forward lever on the tape recorder with the compressed tape, the medical student can skim the lecture much as he would skim a chapter in a book, thus enabling the student to become a participant rather than a spectator.

Another means of independent study that involves compressed speech deals with five 35-minute synchronized slide-tape presentations of material that are highly correlated with course objectives. While regular tapes are available, the compressed version is a boon to the student for not only does he save time but, also, he is forced to attend very carefully.

A third use for compressed speech is in our Automated Message System accessed through dial telephones located in the student laboratories. By dialing a particular number, students have a choice in listening to

---

*Gloria J. Boyle is a research assistant at the School of Medicine, University of Missouri-Columbia, Columbia, Missouri 65201.

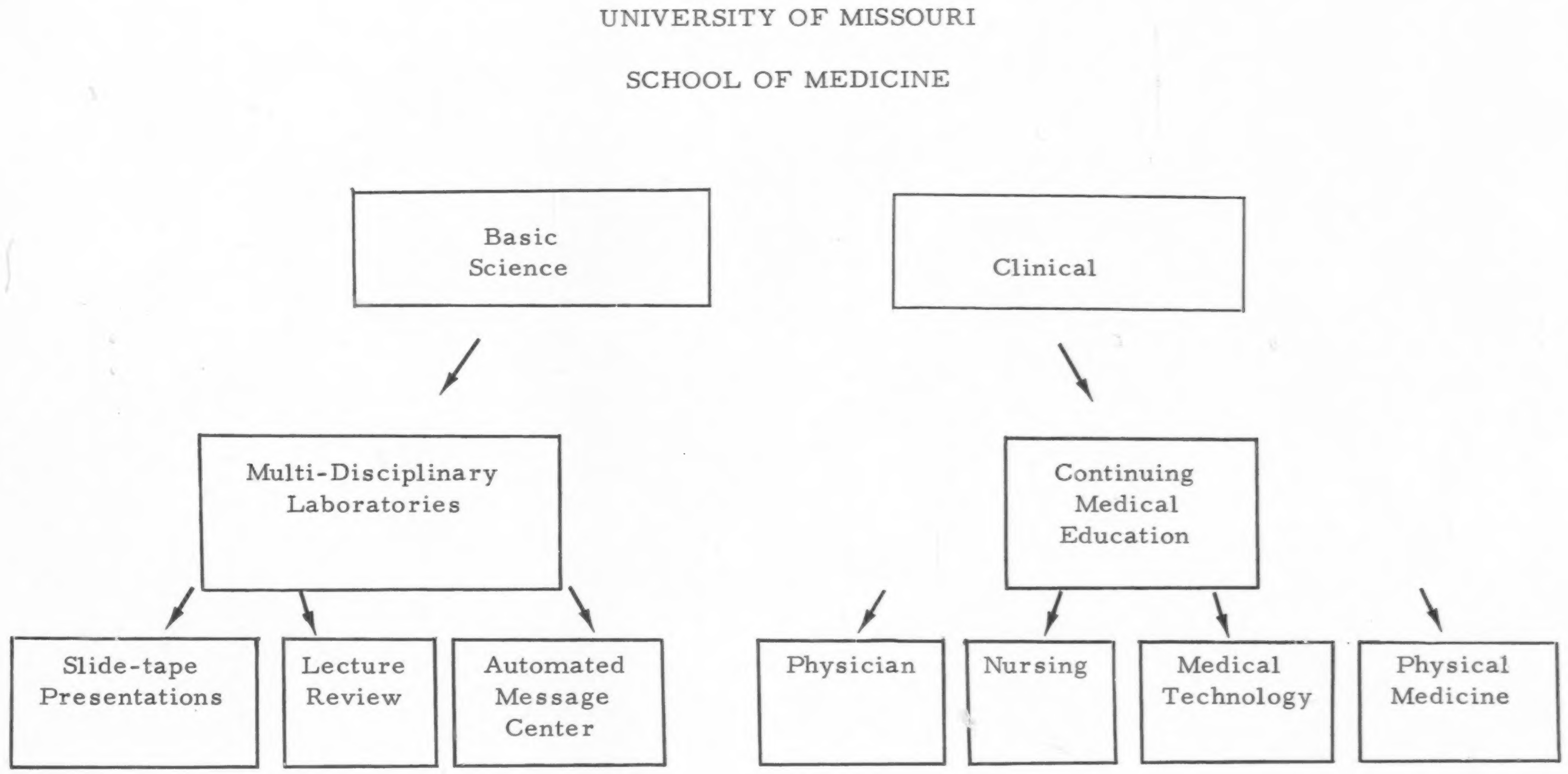

UNIVERSITY OF MISSOURI
SCHOOL OF MEDICINE
Basic Science
Clinical
Multi-Disciplinary Laboratories
Continuing Medical Education
Slide-tape Presentations
Lecture Review
Automated Message Center
Physician
Nursing
Medical Technology
Physical Medicine

Figure 31.1. Compressed speech usage.

lectures, lecture summaries, or other special instructional messages at a slowly paced or compressed rate.

As an extension of the medical school, compressed speech is used to provide further medical education to the personnel of the hospitals (physicians, nurses, medical technologists, and those involved in physical medicine) throughout the state of Missouri. For instance, last year the need for information on the Hong Kong flu was most immediate. By means of a telephone network, 19 hospitals throughout the state had a direct line to the medical school. A physician knowledgeable in the area conducted a lecture followed by questions and answers from the participating hospital faculties. Visual aids were sent out to the hospitals prior to the presentation. A compressed speech copy of the lecture and question-and-answer period was sent to each hospital for review purposes or for absentee staff members. Thus, the visual aids plus the compressed speech copy, at 80% compression, became permanently filed at each hospital providing a constant updating of current clinical material and a quick review for busy hospital personnel.

Production of the compressed copies is accomplished by the Mark II Information Rate Changer (Figure 31.2), various tape recorders, and a Curtin Infonics Tape Duplicator which duplicates three copies simultaneously. With these various applications of compressed speech, we have a built-in opportunity for researching the attitudes and achievement of a well-defined audience using varying compression rates. Through this type of practical research, we may be able to provide a higher degree of individual instruction.

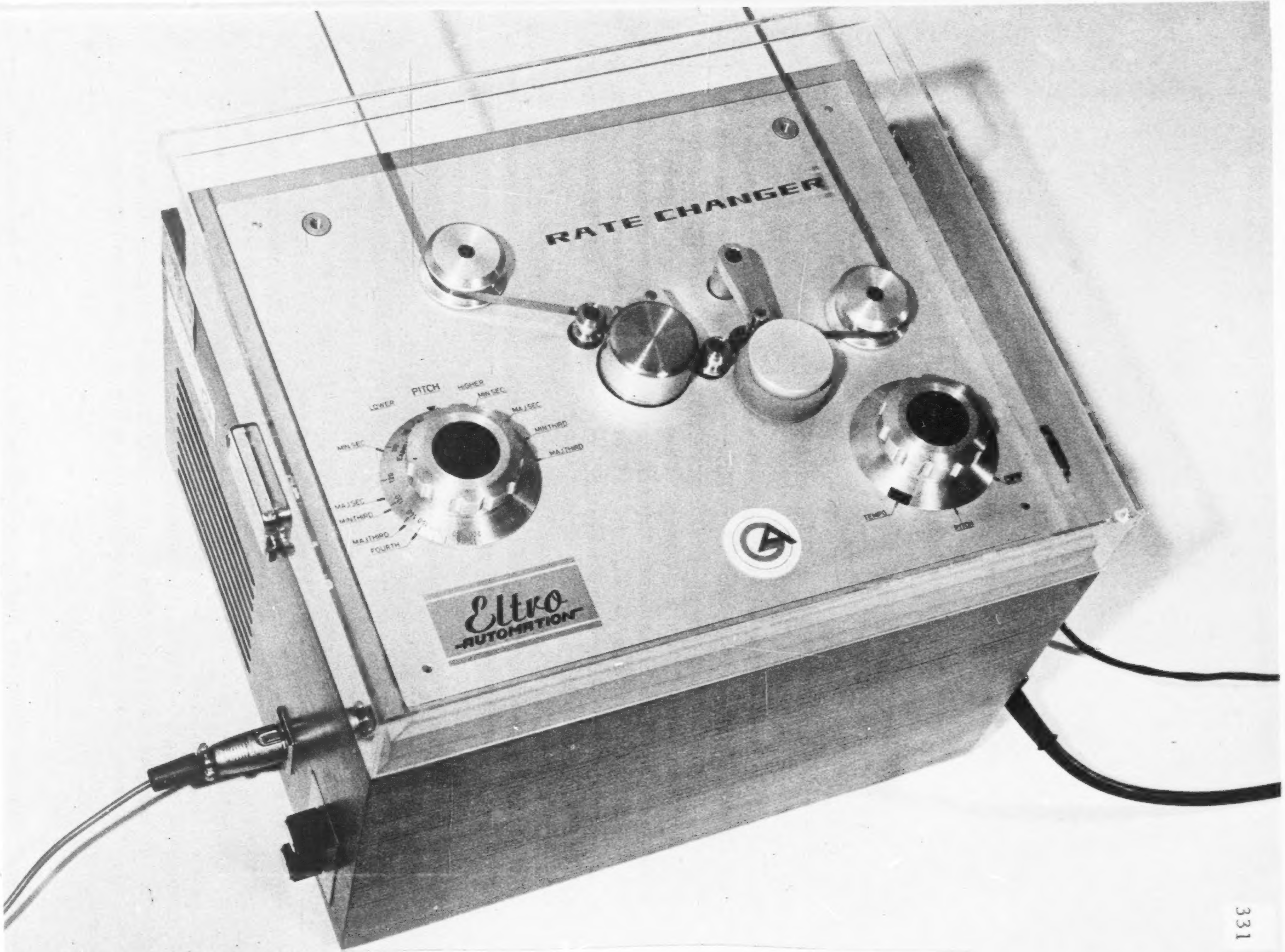

RATE CHANGER
PITCH
LOWER
HIGHER
MIN SEC
MAJ SEC
MIN THIRD
MAJ THIRD
FOURTH
TEMPO
PITCH
OFF
Eltro
AUTOMATION

Figure 31.2. The Mark II Information Rate Changer.

CHAPTER XXXII

THE RELATIONSHIP OF LISTENING SKILLS TO THE UTILIZATION OF COMPRESSED SPEECH

Rolland Callaway, Gerald Gleason, and Barbara Klaeser*

The Problem

Essentially, this pilot study was an attempt to assess the relationship of listening skills to experiences in listening to audio-taped material at varied compressed speeds (rate-controlled recordings). A second purpose of the study was to attempt to assess differences in comprehension of taped material when the sequence of presentation progressed from original taped time (100%) to 55% of original time as compared to a sequence which progressed from 55% of the original taped time to the original time (100%).

This study followed several informal investigations utilizing compressed materials in a graduate class in curriculum planning. The investigator typically uses a number of taped speeches or presentations as a part of the instructional program to supplement the reading requirements and to provide a "common experience" for discussion. The question in the first studies was whether there would be a significant difference in comprehension and understanding if these materials were presented to the students at a compressed rate. Of course, this question was prompted by the reports of investigations which indicated that a loss of comprehension would not take place--in fact, that it might be improved (up to approximately 275-300 words per minute [wpm]). The several informal investigations which were carried out led to somewhat more sophisticated questions such as: (1) what is meant by "comprehension" and "understanding," (2) what effect does level of listening skill have upon the comprehension of compressed materials, and (3) does the utilization of compressed materials affect listening skills--development or deterioration?

---

*Drs. Rolland Callaway and Gerald Gleason are at the University of Wisconsin-Milwaukee. Dr. Callaway is a Professor in the Department of Curriculum and Instruction and Dr. Gleason is Director of Research in the Department of Educational Psychology. Barbara Klaeser is at the Milwaukee School of Engineering, Milwaukee, Wisconsin.

The investigators would like to stress the fact that in the study reported here, compressed materials were utilized and the study conducted in a practical setting as a part of the regular class experiences. There was great interest and discussion concerning the compressed technique on the part of the participants which must be considered in viewing the results and any conclusions drawn. Also, it is important to note that the taped presentations were not specially prepared for the study; thus, there were some problems of fidelity. (The compression was done at the University of Louisville through the cooperation of Dr. Emerson Foulke. The University of Wisconsin-Milwaukee has since acquired an Eltro Automation Rate Changer.)

Research Design and Procedures

The Ss in the study were 40 teachers and school administrators enrolled in a graduate course in curriculum planning during the 1968 summer session at the University of Wisconsin-Milwaukee. The group was divided into three equal groups according to sex and on the basis of pretest scores on the Sequential Tests of Educational Progress (STEP) Listening Test (Form 1A). In a language laboratory setting, the Ss listened to six audio-taped presentations related to the purposes and content of the course during six regular class period (see Table 32.1).

Immediately following the listening session, the Ss took a comprehension test which included 10 simple recall items plus an excerpt of approximately 250 words utilizing the CLOZE (1967) procedure (every fifth word left blank).* The tests were scored, returned to the Ss at the next class session, and the content of the taped presentation served as the topic for class discussion.

During the last week of the class, the STEP Listening Test (Form 1B) was administered as a posttest measure of listening skills.

---

*In a previous study the investigators attempted to devise tests which included items involving more than "simple recall," that is, involving all levels of the cognitive domain of the taxonomy of educational objectives as classified by Bloom (1964) and others. This proved to be most difficult. Continuing with this approach required more time and effort than possible in this study. Thus, the resort to the simple recall and CLOZE (1967) procedure.

TABLE 32.1

PRESENTATION SCHEDULE

Session	Presentation	Compressed Rate % of original taped time 55% to 100% Group A	100% to 55% Group B	100% Group C
1	"Needed: A Unifying Theory of Education" by Harry Broudy	55%	100%	100%
2	"Educational Wastelands or Fertile Fields" by Arthur Bestor and Alan Griffith	65%	85%	100%
3	"The Central Purpose of American Education" by Theodore Brameld	75%	85%	100%
4	"Direction and Redirection for Curriculum Change" by John Goodlad	85%	75%	100%
5	"Sociological Knowledge and Needed Curriculum Research" by Louis Raths	85%	65%	100%
6	"Teaching as Curriculum Decision Making" by Virgil Herrick	100%	55%	100%

Analysis of Data

The first hypothesis was stated as follows:

> $\underline{H}_1$: there will be no indication of significant changes in listening skills of students after listening to a series of materials at compressed rates as compared to students who have listened to the same materials at the original taped rate.

The analysis (using analysis of covariance) indicates that the null hypothesis can be accepted (see Table 32.2). There were no statistically significant changes in the mean scores of the three groups on the STEP Listening Test. However, it is interesting to note that the greatest difference was in the control group which listened to all of the presentations at the original taped rate. The mean score for the group which listened to the material progressing from the original time to 55% changed slightly in a negative direction. There was no change in Group A (55% to 100%).

TABLE 32.2

MEAN SCORES ON STEP LISTENING TEST

	Pretest Mean	Posttest Mean
Group A 55% to 100%	75.8	75.8
Group B 100% to 55%	76.5	74.1 (-2.4)
Group C 100%	76.3	73.4 (-2.9)

While a more detailed analysis of the individual scores is probably in order, we have not done so--primarily because of the time and effort involved and the feeling that further analysis of the data of this study would not lead to further insights. (The individual changes in pre- and posttest STEP Listening Test scores are presented in Table 32.3.)

TABLE 32.3

CHANGES IN INDIVIDUAL PRE- AND POSTTEST SCORES ON STEP LISTENING TEST

No.	Group A 55% to 100%	Group B 100% to 55%	Group C 100%
1	+ 2.8	+ 8.4	-10.5
2	- 6.9	+ 6.9	+12.5
3	- 5.5	- 9.8	+12.5
4	- 6.6	+26.4	-13.9
5	- 2.8	+ 4.3	-13.4
6	- 4.1	-12.5	+ 7.0
7	-13.8	- 1.3	X
8	-12.6	- 4.2	+ 8.4
9	- 2.8	- 2.8	- 8.2
10	- 8.3	-19.4	-13.9
11	+ 4.5	X*	+ 1.4
12	-11.2	- 1.4	- 5.7
13	+ 3.0	X	-13.9
14	+18.1		
	4+, 10-	4+, 7-	5+, 7-

*Did not take both pre- and posttest.

The second and third hypotheses were stated:

$\underline{H}_2$: there will be no significant difference in the comprehension of taped materials when presented at a compressed rate as compared to comprehension of the same materials presented at the original taped time.

$\underline{H}_3$: there will be no significant difference in the comprehension of a series of taped materials presented at rates which progress from 55% of original rate to the original rate as compared to the comprehension of the same series of materials presented at rates which progress from the original rate to 55% of the original rate.

There were no significant differences in comprehension for either the simple recall or the CLOZE (1967) part of the comprehension quiz so Hypothesis_2 and Hypothesis_3 can both be accepted. The mean scores for the three groups on the simple recall quiz (10 items on each quiz) are presented in Table 32.4.

TABLE 32.4

GROUP MEAN SCORES--SIMPLE RECALL QUIZ

Presentation	Group A 55% to 100%	Group B 100% to 55%	Group C 100%
1	5.1	4.9	5.1
2	4.9	4.8	4.4
3	5.8	6.1	5.6
4	4.9	5.3	5.3
5	6.9	6.0	6.9
6	6.1	5.5	6.3

The scores on the CLOZE (1967) procedure part of the quiz are presented in Table 32.5 These percentage figures represent the proportion of the blanks correctly filled in.

Comments

As has been indicated, this is a report of an attempt to study the relationship of listening skills in the practical utilization of rate-controlled recordings. An attempt was also made to assess the effect of progressing from presentations which went from 55% to 100% as contrasted with presentation which progressed from 100% to 55% of original taped time.

While no significant differences either in respect to listening skills or comprehension were identified, the fact that there were no significant differences may be "significant." First, the results seem to indicate that the limited exposure to rate-controlled recordings had little effect on listening skills. Second, the rate of compression seemed to have little effect on comprehension (at least as measured by the instruments used in this study).

TABLE 32.5

GROUP SCORES FOR THE CLOZE PROCEDURE QUIZ

Presentation	Group A 55% to 100%	Group B 100% to 55%	Group C 100%
1	31%	31%	34%
2	64%	62%	62%
3	72%	68%	66%
4	55%	49%	45%
5	68%	64%	68%
6	59%	47%	61%

It is clear (to us, at least) that much research on the listening process and the development of listening skills is essential--especially if "listening" continues to be absolutely necessary for most school learning.

CHAPTER XXXIII

DEAF CHILDREN'S AUDITION OF DISTINCTIVE FEATURES WITHIN FREQUENCY-SHIFTED SPEECH

Daniel Ling*

Problem

Several types of real-time coding amplifiers have been used in attempts to improve the auditory discrimination of Ss with severe high frequency hearing loss. Guttman and van Bergeijk (1958) used the vobanc (Bogart, 1956) to compress the speech spectrum by a factor of two. Johansson (1961) developed a coding amplifier (transposer) in which high frequency speech sounds are heterodyned against a 4,800 Hz reference tone to generate low frequency analog signals. Pimonow (1965), Ling and Druz (1967), Lafon (1967), and Ling and Doehring (1969) each report the use of different types of vocoder (Dudley, 1939). Guttman and Nelson (1968) also describe an instrument for generating a low frequency pulse for every n zero crossings of high frequency fricative sounds in natural speech.

Most recently, Biondi and Biondi (1968) describe a single channel transposer which used a sample and hold process to accomplish multiple transpositions of the input spectrum. The sampling frequency is variable from 1,000-4,000 Hz and the instrument can provide, within the residual hearing range of the S, various degrees of transposition and overlap of the sidebands generated in the sample and hold process. At a high frequency sampling rate, the product of the process is similar to that of Johansson's device.

Without exception, studies employing these various coding amplifiers have shown that deaf Ss are able to learn to discriminate frequency transposed speech. However, results are generally inferior to those obtained with conventional (linear) amplification.

---

*Dr. Daniel Ling is Assistant Professor, School of Human Communication Disorders, McGill University, 3465 Cote des Neiges, Montreal, Quebec, Canada.

Since it is impossible to find completely naive listeners, the problem of comparing coded speech with linearly amplified speech is complicated. Deaf Ss old enough to respond reliably on auditory tests usually have some familiarity with speech in its natural (linearly amplified) form. In contrast, they have no prior experience with transposition which requires the learning of a partly or completely new auditory code.

Bias favoring linear amplification is, therefore, inherent in experiments comparing speech discrimination under the two processes. Two strategies can be adopted to minimize the source of bias. The first is to train Ss under each condition to crude limits of learning (Ferguson, 1956), so that discrimination scores show no significant improvement with further training. This procedure was used by Ling and Doehring (1969) in a programmed learning experiment. The second is to compare Ss' performance on a task which requires no learning or experience at lexical or semantic levels. Such a task was proposed by Travis and Rasmus (1931), who designed a test employing like and unlike pairs of syllables such as /pa-pa/ and /pa-ta/ which demand only a same-different judgment and response.

The purpose of the present experiment was to explore the use of a same-different test paradigm in the evaluation of a frequency transposing device. To determine whether this test strategy yielded similar trends to one involving training to asymptotic performance, the vocoder and Ss previously used by Ling and Doehring (1969) were employed.

Method

Subjects

Ten boys, aged 7-11 years, attending the Montréal Institut des Sourds were selected. All were of average or above average ability. Half of the group had been trained to crude limits of learning with conventional amplification and half with frequency transposition immediately prior to this experiment. All Ss had poorer hearing for high frequency sound than for low. Hearing levels at 1,000 Hz ranged from 70-110 db (mean 87.5 db).

Materials

Stimuli were five sets of three consonants combined with the vowel /a/. The sets were as follows: /s, f, v/; /d, ʒ, z/; /t, d, b/; /f, p, k/; and /n, l, z/. Syllables in each set were combined in all possible ways to yield nine pairs, three of which were the same (e.g., sa-sa) and the remainder different (e.g., sa-fa). Each set was then used to construct five corresponding series each containing 36 same pairs and 36 different pairs. The 72 items in each series were listed in random order and then recorded on tape by a female speaker. The interval between syllables in each pair approximated 0. 25 seconds.

Apparatus

A Uher 5000 tape recorder was used to record and play back the stimuli. The transposing instrument was a vocoder. Described by Ling and Doehring (1969), it analyzed sounds from 1,000 to 4,000 Hz in 10 logarithmically spaced bandwidths. The 10 corresponding analog channels were spaced at intervals of 100 Hz from 100 to 1,000 Hz. The instrument included one linear channel with a frequency range 70-7,000 Hz and switching to permit linear amplification to both ears, transposition to both ears, or linear amplification to one ear and transposition to the other. In the present experiment, the first two of these three conditions were employed. Output of both linear and analog channels to TDH 39 earphones was controlled to approximately 120 db by means of a VU meter.

Subjects responded by pressing one or other of two buttons on a response device. The device, controlled by a logic circuit constructed with Digi-Bits solid state programming modules, incorporated two same- and two different-colored light bulbs. Inaudible pulses recorded on the tape, closed or opened circuits in such a way that when a correct response was made, the appropriate pair of bulbs would light. Thus, automatic feedback to the Ss and examiner was provided on correctness of response.

Procedure

Subjects were first trained to make same and different judgments about pairs of shapes or colors. As soon as satisfactory performance had been obtained, similar pretraining was provided with auditory stimuli under each amplification condition (conventional amplification to both ears and transposition to both ears). The five test series were then administered in a counterbalanced order to each S under each condition of amplification.

Testing took place in a quiet, distraction free room. The S and examiner sat side by side facing the table on which the equipment was arranged. The examiner stopped the tape recorder after each trial, recorded the S's response and restarted the recorder for the next trial. About 10 minutes work was required to complete each series. Two series (one under each amplification condition) were administered on each of 5 consecutive days.

On completion of the testing schedule, Ss were given a series of 100 items in the absence of auditory stimuli and asked to guess whether each item, if heard, would have been the same or different. The purpose of this control procedure was to determine whether Ss tended to respond more frequently to one button than to the other. Bias toward the right was predicted on the basis of work carried out by Bindra, Donderi, and Nishisato (1968).

Results

Data obtained from the control procedure were examined and a significant preference for responses to the right (same) button was found ($t = 3.42$; $p < 0.01$). In all, 53.8% of responses in the absence of sound were made to the right and 46.2% to the left. Results for same and different responses were, therefore, analyzed separately.

The number of pairs correctly identified as being the same and different increased over the testing period. Out of a possible 36, mean same scores increased from 26.3 to 30.4 and mean different scores from 23.0 to 26.1. Nonparametric trend analysis (Ferguson, 1965) showed that both gains were significant beyond the 0.05 level. These increases reflect the extent to which some form of learning occurred.

Subjects' scores under each amplification condition for same and different items correct, pooled across the first five series, are shown in Table 33.1. Differences between Ss, which were significant beyond the 0.01 level (for sames, $F = 13.174$; for differents, $F = 7.893$; both with 9/36 df), were not correlated with pure-tone hearing loss. Conventional amplification was superior to transposition both for pairs which were correctly judged same ($F = 13.182$, df = 1/36, $p < 0.01$) and those correctly judged different ($F = 4.455$, df = 1/36, $p < 0.05$). The series, as expected, were found to be of unequal difficulty (for sames, $F = 4.864$; for differents, $F = 11.316$; both with 4/36 df, $p < 0.01$). In general, series 1 and 4, constructed exclusively with unvoiced consonants, yielded the poorest scores, but there was a significant Subjects x Series interaction for different pairs ($F = 32.833$, df = 4/36, $p < 0.01$). The variance due to Subjects x Conditions interactions was negligible ($F < 0.6$, df = 9/36 for both same and different items correct). Thus, Ss' performance did not reflect previous training under one or other of the two amplification conditions.

Subjects' scores for each same and each different pair were then pooled and the proportion of correct responses obtained under each amplification condition was calculated. Results of this analysis are presented in Table 33.2. Examined within the framework of the distinctive feature system proposed by Chomsky and Halle (1968), these results suggest that pairs differing by more than one distinctive feature are more readily recognized as unlike than those differing by only one. Thus, a smaller proportion of correct responses is associated with /sa-fa/, /ʒa-za/, /da-ba/, and /pa-ka/ than with /da-ʒa/, /fa-ka/, and /na-za/. However, /fa-va/ and /ta-da/, which are differentiated by only one feature, were more frequently judged as unlike than /na-la/ which differs by three. Surprisingly, these comparisons hold good for both conditions of amplification. While transposition changed the frequency characteristics of these sounds, it did not appear to make any of them easier to hear relative to conventional amplification.

TABLE 33.1

THE NUMBER OF PAIRS CORRECTLY JUDGED TO BE SAME OR DIFFERENT BY EACH SUBJECT UNDER EACH CONDITION OF AMPLIFICATION. EACH CELL REPRESENTS RESULTS FOR 350 PRESENTATIONS

	Conventional		Transposition	
Subject	Same	Different	Same	Different
1	151	126	136	109
2	166	115	150	103
3	145	106	128	105
4	153	123	136	106
5	155	142	148	129
6	134	117	127	127
7	120	121	101	132
8	143	107	141	105
9	164	143	157	129
10	156	127	155	128
Sum	1,530	1,227	1,379	1,173

TABLE 33.2

THE PROPORTION OF CORRECT RESPONSES ASSOCIATED WITH EACH PAIR OF SYLLABLES IN RELATION TO THE TWO AMPLIFICATION CONDITIONS EMPLOYED. CONVENTIONAL (CONV.) AND TRANSPOSITION (TRANS.)

Series:							
1.	Item	s-s	f-f	v-v	s-f	s-v	f-v
	Conv.	0.78	0.82	0.79	0.33	0.66	0.88
	Trans.	0.84	0.80	0.69	0.40	0.63	0.84
2.	Item	d-d	ʒ-ʒ	z-z	d-ʒ	d-z	ʒ-z
	Conv.	0.88	0.90	0.77	0.95	0.86	0.42
	Trans.	0.89	0.90	0.82	0.88	0.86	0.45
3.	Item	t-t	d-d	b-b	t-d	t-b	d-b
	Conv.	0.78	0.81	0.82	0.73	0.70	0.62
	Trans.	0.73	0.70	0.74	0.71	0.62	0.53
4.	Item	f-f	p-p	k-k	f-p	f-k	p-k
	Conv.	0.82	0.78	0.87	0.78	0.82	0.45
	Trans.	0.73	0.64	0.72	0.59	0.70	0.45
5.	Item	n-n	l-l	z-z	n-l	n-z	l-z
	Conv.	0.90	0.86	0.77	0.41	0.94	0.82
	Trans.	0.75	0.84	0.69	0.46	0.82	0.80

Discussion

In a previous experiment by Ling and Doehring (1969) using the same *S*s and the same transposition process, no significant differences were found between the two amplification conditions. In the present study, results significantly favored conventional amplification over transposition. In the former, *S*s were trained to asymptotic performance in the discrimination of words through programmed instruction; in the latter, *S*s were not trained to crude limits of learning, and syllables rather than words were employed. Furthermore, scores increased significantly over testing sessions. Had training to crude limits of learning been provided on the task prior to testing, results may not have favored conventional amplification. The findings relating to differences in amplification conditions must, therefore, be regarded as tentative.

Failure to find a significant Subjects x Conditions interaction strongly indicates that Ss previously trained to asymptotic performance on transposition had no particular advantage on this test, which may tap a more fundamental type of discrimination skill than the one involving words. Certainly, the Ss were unable to generalize from their previous experience with transposition.

The use of a test constructed with syllables does not necessarily have the predictive value of a test structured with words. Discrimination of words and word sequences, rather than discrimination of syllables, underlies our every day communication. The successful differentiation of the cognate pairs /t-d/ under conventional amplification, for example, may not have been due to better-than-chance perception of spectral differences. Responses might simply have reflected the extent to which the sound /d/ was audible and the other sound, /t/, inaudible. In a same-different test, audibility versus inaudibility of the releasing phoneme is sufficient to yield a high proportion of correct scores, yet in the discrimination of words, such a contrast might well be meaningless. This is not to say that an adequate test of hearing for speech for deaf Ss cannot be constructed using syllables in a same-different test paradigm. But this example suggests that extremely careful study of stimulus dimensions is required over and beyond the classification of consonants within a distinctive feature system.

The very similar results for transposition and conventional amplification in relation to unvoiced sounds are surprising since the frequency of the speech patterns presented by each are enormously different. That the transposed spectrum falls well within the range audible to each S has been clearly demonstrated with spectrograms of the frequency shifted stimuli (Ling, 1968). Failure to differentiate certain transposed stimuli is, therefore, more likely to be due to discrimination deficits of the type described by Pickett and Martin (1968). These writers have demonstrated that low frequency discrimination of speech-like sounds tends to be poorer among Ss with profound hearing loss than among those with less severe auditory impairment.

The Subjects x Series interaction probably reflects two phenomena: that Ss tended to be consistent though idiosyncratic in responding to stimuli within series and that learning occurred across test sessions.

Differences between series may have been related to the relative audibility of phonemes. On items judged to be different, Scheffé's test showed that results for series 1 and 4 were significantly poorer than those for series 2 and 5. The former were constructed with unvoiced, the latter with voiced consonants. Additionally, series 1 and 4 contained less distinctive feature contrasts than series 2 and 5. Neither possibility, however, satisfactorily accounts for the significant differences between series. For example, the /s-f/ item in series 1 and the /d-b/ item of series 3 are both differentiated on one distinctive feature dimension, namely acute/grave, yet the proportion of correct responses for the voiced pair is significantly greater than for the unvoiced pair. However, consonants within the pairs /ʒ-z/ and /p-k/ are also differentiated by only one feature, namely diffuse/compact, but significantly more correct responses are not associated with /ʒ-z/.

Information theory (Abramson, 1963) would predict a trend for series of items differentiated by several distinctive features to be more frequently judged unlike, than a series differentiated by fewer features. Items which are exceptions to this trend have already been mentioned, and exceptions might simply be due to one or more of the several feature differences being inaudible or to some form of interaction between features. Since a consonant sound may sometimes be identified by modification of the adjacent vowel (Delattre, Liberman, & Cooper, 1955) and the extent of vowel modification depends on the transitions intrinsic to each syllable (Wang & Fillmore, 1961), different results for features in each series might also be expected if different vowels were used.

In brief, considerably more needs to be known about the speech wave correlates of the various features, the simultaneous and successive context effects associated with them, their relative audibility, and the mechanisms by which deaf Ss encode speech, before explanation of differences found between series in this study rises above a conjectural level.

Conclusions

The use of pairs of like and unlike syllables which Ss judge to be the same or different was shown to have some disadvantages. Subjects could use related cues (such as audibility versus inaudibility) rather than spectral differences in making discriminations.

Improvement in scores over sessions showed that training to crude limits of learning with same-different test material would be necessary if valid comparisons of amplification conditions are to be made on the basis of data yielded by such a test.

Results suggested that discrimination of distinctive features was dependent upon context.

Results differed from those obtained with words in a previous study in that present findings tentatively suggest the superiority of conventional amplification over this form of transposition.

The hypothesis that Ss previously trained to asymptotic performance with words under one or other amplification condition would achieve better results under that condition when tested with like and unlike pairs of syllables was not supported.

The development of an adequate test of discrimination for deaf Ss which uses a same-different paradigm requires further research.

ACKNOWLEDGMENTS

This research was supported by Canadian Federal Provincial Health Grant 604-7-573 to the late Dr. Hollie E. McHugh. The author is grateful to Frère Viateur Gervais of the Montréal Institut des Sourds for providing space and subjects, to the Zenith Corporation for the loan of the transposing device, and to Miss Janet Ashdown for the data collection.

CHAPTER XXXIV

THE EFFECTS OF TRAINING ON THE INTELLIGIBILITY AND COMPREHENSION OF FREQUENCY-SHIFTED TIME-COMPRESSED SPEECH BY THE BLIND*

Paul E. Resta**

Introduction

Recently, attempts have been made to compress speech in time through the use of mechanical and electronic techniques. The speech which is speeded in this manner is called time-compressed speech. The techniques vary widely, however, in their complexity, sophistication, expense, and current availability. One of the more sophisticated techniques is the sampling method of compressing speech. Words which are speeded using the sampling technique are easily understood and free of pitch distortion, but the necessary equipment is expensive and not readily available. It is anticipated that a time lag of several years will exist before the sampling technique equipment and materials will become widely available to blind learners.

In contrast to the sampling method, the simplest, least expensive, and most widely available technique for compressing speech has received little attention by researchers. This technique, known as the speed changing method, involves playing back a recording at a faster speed than that at which it was recorded. The rapid playback conveniently accelerates the speech rate but also results in a frequency shifting of the original speech sounds. Nevertheless, a number of blind listeners who are unable to accomplish their required reading in the time available have trained themselves to listen to their tapes and records played back at double speed with little loss of comprehension (Taylor, 1967).

---

*A complete report of this research project is contained in the author's "The Effects of Training on the Intelligibility and Comprehension of Frequency-Shifted Time-Compressed Speech by the Blind," AEEM-298, Goodyear Aerospace Corporation, Litchfield Park, Arizona, May 15, 1968.

**Dr. Paul E. Resta is Group Head, Instructional Management Systems, Southwest Regional Laboratory for Educational Research & Development, 11300 La Cienega Boulevard, Inglewood, California 90304.

Dependent Variables

Improving the intelligibility and comprehension of frequency-shifted (FS) time-compressed speech by training the listener is an obvious possiblity but one which has, as yet, remained unexplored. The objectives of the present study were to determine whether training could significantly increase the intelligibility and comprehension of FS time-compressed speech by blind students and to investigate the effects of selected permutations of four potentially relevant variables in a training situation. The four variables selected for the study included practice listening; speech rate presentation mode; type of training material; and feedback.

Practice Listening to FS Speeded Speech

Since differences were found between groups provided with practice listening to nonfrequency-shifted (NFS) speeded speech and those which had not, it was similarly hypothesized that Ss provided with practice listening to FS compressed speech would show higher performance than those who had not.

Speech Rate Presentation Mode

Based on the findings of NFS compressed speech research, it was hypothesized that a gradually increasing speech rate would result in greater intelligibility and comprehension of FS compressed speech than would a constant speeded rate. The gradual increasing of speech rate may be considered to be a means of successively approximating a difficult task. Successive approximation has long been established as an effective technique for learning many complex tasks and its superiority over direct practice of the terminal task has been demonstrated in many instances.

Type of Training Material

In listening to FS speeded speech the listener is not only confronted with the problem of receiving information rapidly, but also with accepting this increased information flow in terms of FS sound components. Discriminating the FS speech may be similar in some respects to translating a foreign language. The learner has to associate new sounds (different both in terms of frequency and syllable durations) with their unspeeded counterparts. A potentially effective training program would appear to be one which combines two strategies: (1) the organization of verbal stimuli (according to their phonemic similarities and contrasts) to facilitate the discrimination of phonemes at the speeded level of presentation; and (2) the separation of the problem of the discrimination of the individual FS message units from that of the problem of increased information flow (at least in the initial stages of training).

The materials developed in the linguistic approach to teaching reading (e.g., Bloomfield & Barnhart, 1961) are consistent with both of the two strategies for they are based on: (1) a careful organization and sequencing of verbal stimuli; and (2) the similarity and contrast of the speech sound components. Typically, the linguistic materials first present short lists of individual words followed by the presentation of these same words in reading passages. This arrangement allows the Ss to discriminate selected individual FS message units before they are imbedded into passages of continuous discourse.

Another desirable feature of the linguistic approach is that it uses a logical "building-block" approach in the development of sequences of verbal stimuli. This approach provides greater opportunity for practice of previously learned discriminations throughout the training program. Based on the above considerations, it was decided that the research project should include a comparison of the linguistically structured verbal materials and the narrative, explanatory continuous discourse materials used in previous time-compressed speech research.

Feedback

The use of feedback has not been explored in previous time-compressed speech training research. The results of a vast array of laboratory and classroom experiments, however, clearly indicate that feedback is an important condition for effective learning (Travers, 1966). It was thus decided to investigate the effects of feedback on the learning performance of Ss exposed to FS time-compressed speech.

Hypotheses

A directional hypothesis was presented based on the following four research hypotheses:

H_1: practice listening to speeded speech would result in higher performance than no practice listening to the speeded stimuli.

H_2: training with linguistic materials would result in more effective discrimination of FS time-compressed speech than conventional continuous discourse materials.

H_3: a gradually increasing speech rate would result in higher performance than a constant speeded rate which, in turn, would be more effective than a constant unspeeded rate.

H_4: feedback would result in higher performance than a no-feedback treatment condition.

Description of Treatments

In the present study seven distinct treatment conditions were developed to assess the training effectiveness of specific interactions of the following variables: practice listening to FS time-compressed speech; type of training material; speech rate presentation mode; and feedback.

Practice Listening to FS Time-Compressed Speech

The two treatment conditions related to this variable consisted of (a) providing Ss in Treatment Groups 1-6 with practice listening (or exposure) to FS time-compressed speech; (b) providing Ss in Treatment Group 7 with practice listening to only unspeeded listening materials. The amount of listening practice to FS compressed speech provided the treatment groups is shown in Table 34.1. Groups 1, 3, and 5 (gradually increasing rate groups) had 2.31 hours of listening practice; Groups 2, 4, and 6 (constant speeded groups) had 2 hours of listening practice; and Group 7 had no practice listening to the speeded verbal stimuli. Table 34.2 indicates the training session times for all treatment groups. The session time differentials between Tables 34.1 and 34.2 represent the unspeeded materials presentation time. The unspeeded material consisted of instructions, test questions and alternative choices, and feedback (for Groups 1 and 2).

TABLE 34.1

COMPRESSED SPEECH LISTENING PRACTICE TIME (IN MINUTES) FOR ALL TREATMENT GROUPS

Treatment	Training session						Total
Group	1	2	3	4	5	6	
1	30	26	22	21	20	20	139
2	20	20	20	20	20	20	120
3	30	26	22	21	20	20	139
4	20	20	20	20	20	20	120
5	30	26	22	21	20	20	139
6	20	20	20	20	20	20	120
7	0	0	0	0	0	0	0

TABLE 34.2

TRAINING SESSION TIME (IN MINUTES) FOR ALL TREATMENT GROUPS

Treatment Group	Training session 1	2	3	4	5	6	Total
1	50	45	42	41	40	40	258
2	50	40	40	40	40	40	250
3	40	28	26	26	25	25	170
4	35	25	25	25	25	25	155
5	45	35	30	30	30	30	200
6	40	30	30	30	30	30	190
7	60	50	50	50	50	50	310

Type of Training Material

The two treatment conditions related to this variable consisted of linguistic training and continuous discourse training. In the linguistic training treatment condition, the verbal stimuli were organized and sequenced according to their linguistic components. Individual words were presented singly, followed by presentation of words in short passages of continuous discourse, the length of which was progressively increased throughout training sessions.

In the continuous discourse training treatment condition, the reading selections varied from 4 to 15 paragraphs in length and were similar to the type of material used in previous NFS compressed speech research (Voor, 1962).

Speech Rate Presentation Mode

Three speech rate presentation modes were utilized in this experiment including: (a) a constant normal speech rate; (b) a gradually increasing speech rate; and (c) a constant speeded speech rate. In the constant normal speech rate treatment condition, all of the stimulus materials were presented at the criterion oral reading rate of 244 syllables per minute (spm). As shown in Figure 34.1, the terminal speech rates for the gradually increasing speech rate mode training sessions were 380, 430, 460, 475, and 488 spm, respectively. In the constant speeded speech rate presentation mode, all of the verbal stimulus materials were presented at the criterion speeded speech rate of 488 spm.

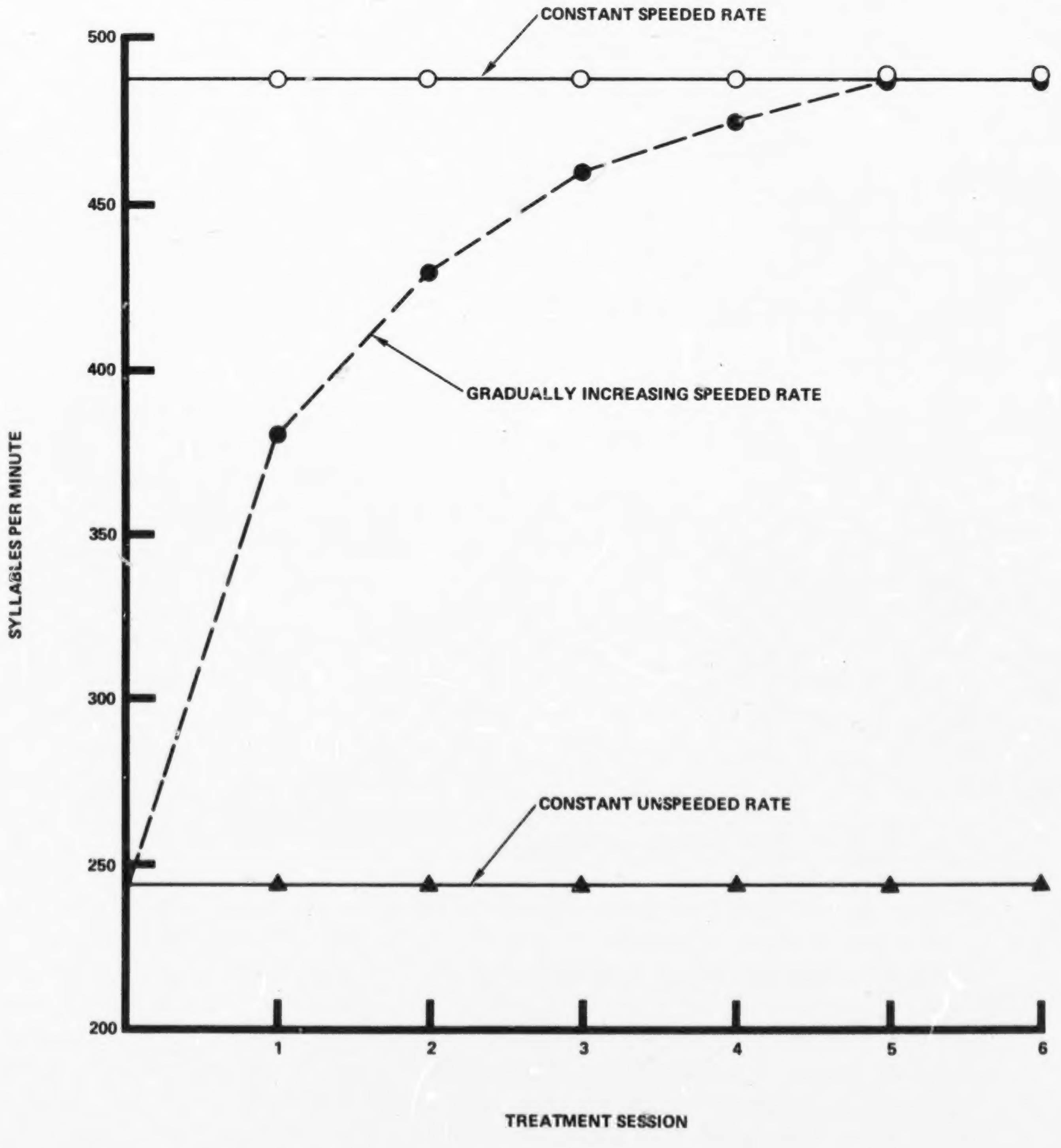


Figure 34.1. Treatment session speech rates for the constant speeded, gradually increasing, and constant unspeeded speech rate presentation modes.

Feedback

Two treatment conditions were related to this variable. In both treatment conditions the Ss heard single words and short sentences which were presented at a speeded rate. They were then asked to identify the message by saying it aloud softly. In the feedback condition, following a short pause, the Ss were presented with the verbal stimuli repeated at an unspeeded rate. In the no-feedback treatment condition repetition of the speeded verbal stimuli (either in speeded or unspeeded form) was not provided the Ss.

Subjects

The Ss consisted of 70 blind, institutionalized students of both sexes (grades 7-12) from the Ohio State School for the Blind. The Ss were native speakers of English. None had any previous exposure to speeded speech or listening instruction programs, and all were without hearing impairment. The Ss were blocked by grade and randomly assigned to treatment groups using the table of random numbers. No attempt was made to classify by sex or visual status, as the research literature does not support the notion of differential compressed speech listening abilities related to these categories.

Experimental Design

A posttest-only control group design was used in which the independent variables were listening practice to FS time-compressed speech, type of training, speech rate presentation mode, and feedback. Intelligibility (as measured by the Black Multiple-Choice Intelligibility Test, Form C) and listening comprehension (as measured by the Sequential Tests of Educational Progress Listening Subtest, Form 2A) were the dependent variables. A single classification analysis of variance was used in which comparisons were made between permutations of the independent variables. The seven distinct permutations of variables selected for the study included, can be seen in Table 34.3.

It is obvious from an inspection of Table 34.3 that only a limited number of the possible permutations were included in the present study. Other permutations of interest (e.g., ACEF, AGEF, BG) had to be excluded because of the limited number of Ss. The ABCE-ABC and ABDE-ABD permutations were selected to provide a comparison of the effects of feedback with type of training material (linguistic), speech rate presentation mode, and listening practice to FS time-compressed speech held constant. The ABC-ACF and ABD-ADF permutations were selected to provide a comparison of the effects of type of training material with listening practice to FS compressed speed and speech rate presentation mode held constant. The ACF and ADF-FG permutations were selected to provide a comparison of the effects of listening practice to FS time-compressed speech with type of material held constant. The ABCE-ABDE, ABC-ABD, ACF-ADF permutations were selected to provide a comparison on the effects of speech rate presentation mode holding type of training material and feedback constant.

TABLE 34. 3

SCHEMATIC REPRESENTATION OF EXPERIMENTAL DESIGN

Treatment group	Treatment description*
1	ABCE
2	ABDE
3	ABC
4	ABD
5	ACF
6	ADF
7	FG

* A = Listening practice to FS time-compressed speech
B = Linguistic training material
C = Gradually increasing speech rate
D = Constant speeded speech rate
E = Feedback
F = Continuous discourse material
G = Listening practice only to unspeeded speech
(constant unspeeded speech rate)

Procedures

The experiment proper was conducted during 9 consecutive school days (excluding weekends). Each treatment group received six training sessions and two testing sessions.

Training Phase

At the beginning of the first training session, the Ss were directed to the listening stations. The Ss were then instructed on the location and operation of the station headsets.

Following a presentation of a brief "warm-up" listening selection, the Ss were queried about their ability to hear the recording and the comfort of the headsets. Adjustments were made as required followed by a presentation of the initial instructions at an unspeeded rate.

After the instructions were given, any questions regarding the nature of the task to be performed by the Ss were answered by the E. Following this the training stimuli were presented.

Treatment Group 1 was presented with stimulus materials arranged according to their linguistic components. The verbal stimuli were presented according to the following scheme:

Speeded Stimuli		Unspeeded Feedback
"can"	------2-second pause------	"can"
"Dan"		"Dan"
"Dan ran a van."		"Dan ran a van."
"Dan ran a tan van."		"Dan ran a tan van."

The students received the following instructions: "You are going to hear some words that are presented at a fast speed. Listen carefully and say aloud softly what you think each fast word is. After a short pause, you will hear the same word at a normal speed. If you are not sure what the fast word is, try to guess. Ready--------the first word is--------can (speeded)--------(2 seconds)--------can (unspeeded)----------man (speeded)--------(2 seconds)--------man (unspeeded)--------etc."

Similar instructions were used for the short sentences that followed the individual words. Students were not requested to say aloud any verbal stimuli longer than a sentence of 10 words. In the first session the stimulus materials were presented initially at a rate of 244 spm and gradually increased to a rate of 380 spm. Figure 34.1 shows the initial and terminal speech rates for each session.

Treatment Group 2 was subjected to the same treatment conditions as Treatment Group 1 with the exception that the speech rate was held constant at the criterion speeded rate (488 spm) during all six sessions.

Treatment Group 3 received the same training conditions as Treatment Group 1 with the exception that no feedback was provided.

Treatment Group 4 received the same training as Treatment Group 2 except that no feedback was provided.

Treatment Group 5 was presented with continuous discourse textual materials at a rate which was gradually increased to the criterion speeded rate during the first five sessions.

Stimulus materials consisted of reading passages of varying length (ranging from approximately 4 to 15 paragraphs). The following instructions were provided the Ss: "You are going to hear a reading selection presented at a fast speed. Listen carefully and see how much of it you can understand."

Treatment Group 6 was presented with the same materials as Treatment Group 5, but the speech rate was gradually increased to the criterion speeded rate over the first five sessions.

Treatment Group 7 was presented with the same stimulus materials as Treatment Groups 5 and 6, but they were presented at a constant unspeeded rate (244 spm).

Testing Phase

Separate test sessions were required for the intelligibility and the comprehension tests.

Intelligibility test. The intelligibility test consisted of three word lists from the Black Multiple-Choice Intelligibility Test. The test item number was presented at a normal rate followed by presentation of the speeded word after a 1-second pause. The S was then provided with four word choices (including the stimulus word) at the unspeeded rate. Subjects indicated their choice of alternatives by marking the appropriate box on their braille cell answer sheets.

Comprehension test. The reading selections were presented to the Ss at the rate of 488 spm. Each listening selection was then followed by its multiple-choice questions presented at a normal rate. The Ss indicated their choices of alternatives by marking on braille cell answer sheets.

Experimental Equipment

All stimulus materials were presented to the Ss via the Dukane Triumph 60 solid state Level II Audio Learning Lab. This lab consisted of a central console and 10 listening station booths, aligned in two rows on opposite sides of the room. The console tape deck consisted of an Ampex dual capstan drive tape recorder.

Results

Intelligibility

The mean scores of treatment groups, as shown in Table 34.4, ranged from a high of 58.7 (Group 1) to a low of 44.0 (Group 7) on the 81-item test. The standard deviation of the scores varied from 7.13 (Treatment Group 4) to 13.81 (Treatment Group 7). A comparison of the treatment group means ordered according to the hypotheses, reveals that the relative magnitude of the means occurred in the predicted order with the exception of the Treatment Groups 5 and 6 reversal.

TABLE 34.4

INTELLIGIBILITY TEST SCORES FOR ALL TREATMENT GROUPS

Treatment group	Mean	S.D.
1	58.7	7.87
2	55.8	7.13
3	52.6	7.96
4	49.2	9.61
5	49.2	9.61
6	50.6	9.25
7	44.0	13.81

As shown by the histogram in Figure 34.2:

1. All groups which were provided with practice listening to FS time-compressed speech (Treatment Groups 1-6) had higher mean scores than did the group which had no practice with the speeded stimuli (Treatment Group 7).
2. All groups which used the linguistic training materials (Treatment Groups 1-4) had higher means than those obtained by the groups using the continuous discourse training materials (Treatment Groups 5-7).
3. The linguistic training groups receiving feedback (Treatment Groups 1 and 2) obtained higher mean values than did the no-feedback groups (Treatment Groups 3 and 4).
4. The linguistic training, gradually increasing speech rate groups (Treatment Groups 1 and 3) obtained higher mean values than did the linguistic training, constant speeded speech rate groups (Treatment Groups 2 and 4). However, a reverse relationship was observed in the continuous discourse training condition. The constant speeded speech rate group (Treatment Group 6) obtained a higher mean intelligibility

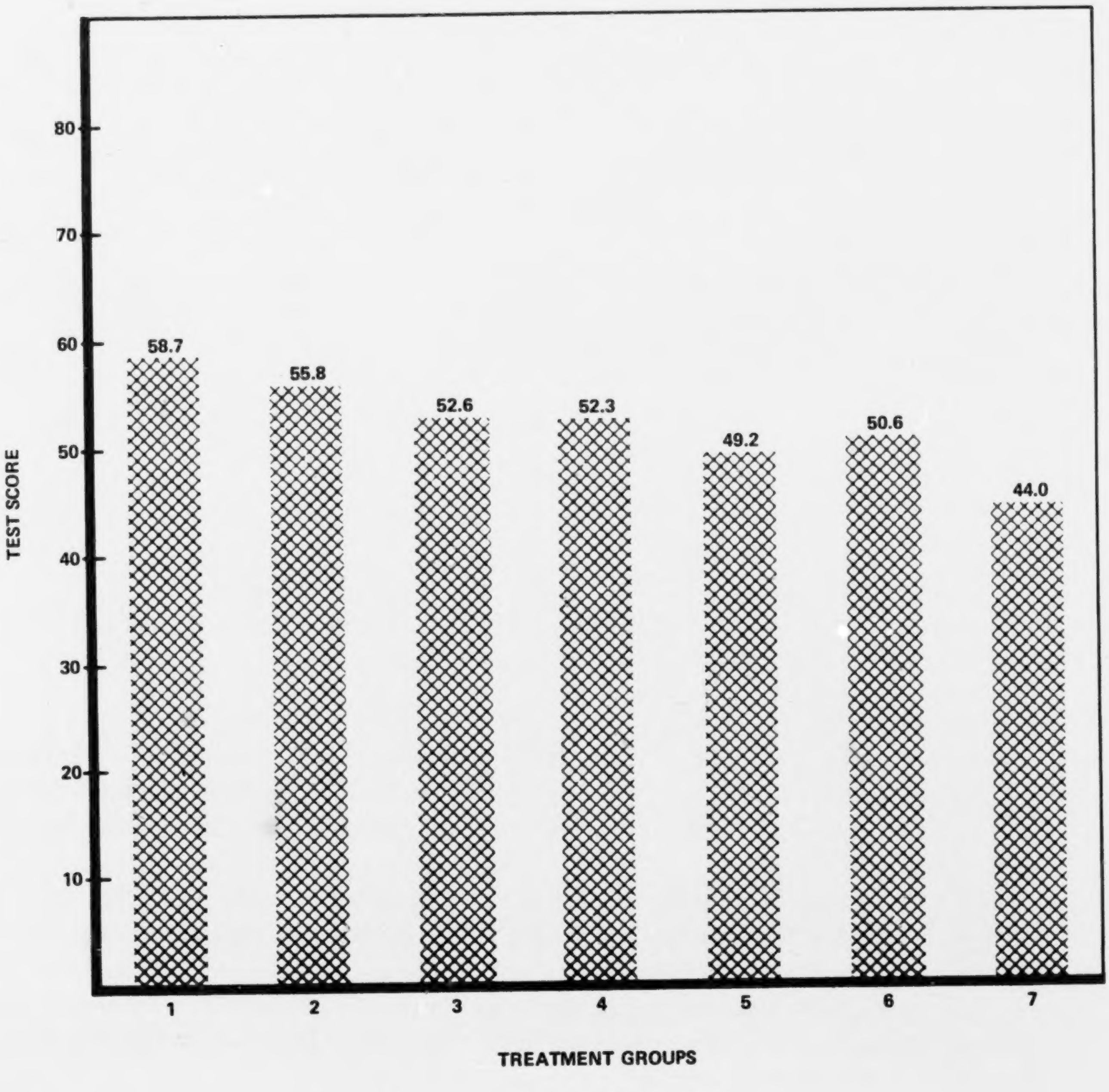


Figure 34. 2. Mean intelligibility test scores of ordered treatment groups.

score than did the gradually increasing and the constant unspeeded groups (Treatment Groups 5 and 7).

Table 34.5 shows the ANOVA table for a single classification analysis of variance of the intelligibility criterion test scores. The obtained $\underline{F}$ value of 2.34 is significant beyond the 0.05 level of confidence. Using the Newman-Keuls method for testing the difference between all pairs of means, the only mean difference found to be statistically significant is the one between Treatment Group 1 (practice listening with FS speeded speech, linguistic training materials, gradually increasing speech rate, and feedback) and Treatment Group 7 (practice listening to unspeeded speech, continuous discourse training materials, constant unspeeded speech rate, and no feedback).

TABLE 34.5

ANALYSIS OF VARIANCE OF INTELLIGIBILITY TEST SCORES FOR ALL GROUPS

Source of variance				
Treatment	6	1,334.89	222.48	2.34*
Within groups	63	5,966.20	94.70	-
Total	69	7,301.09		

*$p < 0.05$

	Treatment
Newman-Keuls	$1\underline{23456}7$

Comprehension

The mean percentage scores of treatment groups, as shown in Table 34.6, ranged from a high of 44.0 to a low of 34.0 on the 80-item test. The magnitude of the standard deviations varied from 7.13 to 11.86.

A comparison of the treatment group means, ordered according to the directional hypotheses, indicates that the means of Treatment Groups 3, 4, and 5 are not consistent with the predicted sequence. The table also shows that:

1. All groups provided with practice listening to the FS time-compressed speech (Treatment Groups 1-6) obtained higher means than the group which had no practice with speeded speech (Treatment Group 7).
2. No gross differences can be observed between the linguistic training

and continuous discourse groups.

3. The linguistic training groups receiving feedback (Treatment Groups 1 and 2) obtained higher means than did the no-feedback linguistic training groups (Treatment Groups 3 and 4).

4. Two of the three gradually increasing speech rate groups (Treatment Groups 1 and 5) had higher mean scores than did the comparable groups having a constant speeded speech rate presentation mode (Treatment Groups 2 and 6). A reverse relationship was found in Treatment Groups 3 and 4. In this instance the linguistic no-feedback constant speeded treatment group obtained a slightly higher mean (0.3 difference between means) than the comparable gradually increasing speech rate group.

TABLE 34.6

COMPREHENSION TEST SCORES FOR ALL TREATMENT GROUPS

Treatment group	Mean	S.D.
1	44.0	7.13
2	42.7	11.12
3	38.9	10.43
4	39.2	7.29
5	42.9	8.14
6	38.0	8.47
7	34.0	11.86

A single classification analysis of variance of the STEP Listening Comprehension Test scores was performed. As shown in Table 34.7, the overall difference among the means tested by analysis is not significant.

TABLE 34.7

ANALYSIS OF VARIANCE OF STEP LISTENING SUBTEST SCORES FOR ALL TREATMENT GROUPS

Source of variation	df	SS	MS	F
Treatment	6	735.38	122.56	1.38
Within groups	63	5,561.50	88.28	-
Total	69	6,296.88		

Discussion

The major finding of the present study was that a discrimination training procedure incorporating practice listening to FS time-compressed speech; linguistic training materials; gradually increasing speech rate; and feedback can significantly increase the intelligibility of FS time-compressed speech. This finding is consistent with the findings of Staats, Staats, and Schutz (1962) and other investigators that stimulus discrimination pretraining results in positive transfer to a later task when the same or similar stimuli are presented in that task. Although a statistically significant difference in intelligibility performance was only found between Treatment Group 1 (group receiving all hypothesized optimal training conditions) and Treatment Group 7 (control group), the trend of the mean values of the other treatment groups is provocative. With one exception, the magnitude of the intelligibility test means from Treatment Groups 1-7 is generally consistent with the predicted order.

The intelligibility findings have important implications for future compressed speech research. The major limitation of the widely available speed-changing method has been its reported initial low intelligibility as compared to speech compressed by the sampling method. This limitation has been the primary rationale for not exploring the application of the speed-changing method to the information acquisition problems of the blind student. It is hoped that the finding that a modest amount of training can significantly increase the intelligibility of FS compressed speech will result in the speed-changing method receiving greater research attention in the future.

In contrast to intelligibility performance, no statistically significant differences were found between groups on the comprehension measure. The 10-point differential between Groups 1 and 7, however, is provocative. Two possible explanations for these findings are as follows:

1. It is well established that perception of the individual message units is a necessary but not sufficient condition for comprehension of continuous discourse. Although intelligibility was improved through training, it is possible that a higher percentage of word intelligibility is required for significant improvement in comprehension than was obtained in the present study. It is also possible that, with more extensive training than the brief amount provided in the present study, greater gains would be made in both intelligibility and comprehension performance.
2. The lack of a concomitant statistically significant increase in comprehension performance may at least partially be a function of the rate at which the information was presented. The findings of a recent study using NFS time-compressed speech by Foulke and Bixler (1967) indicate that a marked loss in comprehension, without appreciable loss of intelligibility, occurs at speech rates exceeding 325 words per minute (wpm). It can be hypothesized, therefore, that the 350 wpm speech rate used in

this study (assuming similar syllable/word ratios in the Foulke and Bixler materials and those used in this investigation) may have exceeded the information processing capabilities of many of the Ss. To investigate this possibility it is suggested that the effects of training on the intelligibility and comprehension be studied at speech rates at or below the 325 wpm rate.

Because of the restricted N, there were a number of permutations of variables that were not included but would be worthy of investigation in future studies. For example, the use of feedback with continuous discourse training materials may also merit exploration, not only in FS time-compressed speech research, but in NFS time-compressed speech as well. In addition other dimensions of the four independent variables identified in the present study should be investigated in future compressed speech research, e.g., greater amounts of listening practice to FS time-compressed speech, shorter or longer acceleration intervals in gradually increasing the speech rate, and varied amounts of feedback. Greater research attention should also be directed to the nature of the training materials themselves. Little has been done, thus far, in identifying and analyzing the variables associated with the stimulus materials other than the mechanical aspects of accelerating the speech rate. Little is known, for example, about the effects of such variables as passage length, word difficulty, and type of subject matter on the intelligibility and comprehension of time-compressed speech. Typically these variables have either been ignored or assumed to be controlled through the use of formulas designed for measuring the difficulty level of reading materials. Whether the formulas widely used in compressed speech research (e.g., the Dale-Chall formula) are valid measures of the difficulty of listening materials has not as yet been established.

Another recommendation, based on the results of the present study, is that greater attention should be focused on the dependent variables used in compressed speech research. Intelligibility and comprehension have been the only dependent variables included in previous studies of speeded speech. The intelligibility tests have typically consisted of the recognition of single words, while comprehension has been measured by asking questions regarding the content of continuous discourse passages ranging in length from a few paragraphs to several pages. Between these two types of measures there is obviously a large "no-man's-land" about which we know very little. It is suggested that in future compressed speech research, measures such as the recognition/recall of short sentences, long sentences, and short paragraphs also be used to provide more definitive information on this uncharted territory. Measures such as the above may help provide the badly needed data on the interrelationships between the quantity of information and the rate at which it is transmitted.

LIST OF REFERENCES

Aaronson, D. Immediate recall of compressed speech. Unpublished manuscript, Harvard University, Center for Cognitive Studies, 1966.

Aaronson, D. Temporal factors in perception and short-term memory. Psychological Bulletin, 1967, 67, 130-144.

Aaronson, D. Temporal course of perception in an immediate recall task. Journal of Experimental Psychology, 1968, 7, 129-140.

Aaronson, D., & Markowitz, N. Immediate recall of normal and compressed auditory sequences. Paper presented at the Eastern Psychological Association, Boston, April 1967.

Aaronson, D., & Steinberg, S. Effects of presentation rate and signal-to-noise ratio on immediate recall. Paper presented at the Eastern Psychological Association, Philadelphia, April 1964.

Abramson, N. Information theory and coding. New York: McGraw-Hill, 1963.

Agnello, J. G. A study of intra- and inter-phrasal pauses and their relationship to the rate of speech. Unpublished doctoral dissertation, Ohio State University, 1963.

Ahlstrom, K., Risberg, A., & Lindle, V. Experiment in learning to discriminate frequency-transposed speech. (Report No. 36) Institute of Education, Uppsala University, Uppsala, Sweden, 1968.

Ahmend, R., & Fatechand, R. Effects of sample duration on the articulation of sounds in normal and clipped speech. Journal of the Acoustical Society of America, 1958, 31, 1022-1029.

Allen, H. S., Jr. The Eltro Information Rate Changer Mark II: Simple quality speech compression. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967.

Altshuler, M. W. Responses to expanded speech by hard of hearing aged subjects. (Unpublished doctoral dissertation, Boston University, 1963) Dissertation Abstracts, 1965, 25, 7069.

Asher, J. J. An experimental study of listener comprehension of news commentary when rate of delivery is varied. Unpublished master's thesis, University of Houston, 1955.

Asher. J. W. Intelligibility tests: A review of their standardization, some experiments, and a new test. Speech Monographs, 1958, 25, 14-28.

Attneave, F. Some informational aspects of visual form perception. Psychological Review, 1954, 61, 183-193.

Atwood, J., Babcock, M., Gazdag, J., Peterson, L., & Schill, J. Speech analysis and classification. Biological Computer Laboratory Report No. 67.2. Champaign, Ill.: University of Illinois, Biological Computer Laboratory, 1967.

Axelrod, S. Underestimation of dichotic click rates: Results using methods of absolute estimation and constant stimuli. Psychonomic Science, 1968, 12, 133-134.

Babcock, M. Speech characterization, i.e., analysis, synthesis, and recognition. Biological Computer Laboratory Report No. 68.2. Champaign, Ill.: University of Illinois, Biological Computer Laboratory, 1968.

Baker, H. J., & Leland, B. Detroit tests of learning aptitude, examiner's handbook. Indianapolis: Bobbs-Merrill, 1967.

Ball, J. H. Synthesis of original vocal pitch in accelerated-playback. Unpublished master's thesis, Massachusetts Institute of Technology, 1961.

Barabasz, A. A study of recall and retention of accelerated lecture presentation. Journal of Communication, 1966, 18, 283-287.

Beasley, D. Perception of time-varied synthetic sentences. Unpublished doctoral dissertation, University of Illinois, 1969.

Beasley, D., & Zemlin, W. Time and frequency-altered speech signals: A bibliography. University of Illinois, 1969. (Mimeo)

Beasley, D., Zemlin, W., & Silverman, F. Influence of frequency transposition on judgments of intelligibility and sex. Paper presented at the American Speech and Hearing Association Convention, Chicago, 1969.

Beetle, D., & Chapman, W. Flexible analog time compression of short utterances. Institute of Electrical and Electronic Engineers Transactions on Audio and Electro-acoustics, 1968, AU-15, 12-20.

Beighley, K. C. An experimental study of the effect of four speech variables on listener comprehension. (Unpublished doctoral dissertation, Ohio State University, 1958) Dissertation Abstracts, 1958, 18, 1894-1896.

Beil, P. Frequency analysis of vowels produced in helium-rich atmosphere. Journal of the Acoustical Society of America, 1967, 34, 347-349.

Bellamy, M. J. An experimental study to compare the comprehension of speeded speech by blind and sighted children. Unpublished master's thesis, University of Texas, 1966.

Bellefleur, P. The effect of an alternation in the time-frequency characteristics of selected speech samples on pathological ears. Unpublished doctoral dissertation, Wayne State University, 1964.

Bennett, D., & Byers, V. Increased intelligibility in the hypacusic by slow-play frequency transposition. Journal of Auditory Research, 1967, 7, 107-119.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. Differential aptitude tests. New York: The Psychological Corporation, 1961.

Bindra, D., Donderi, D. C., & Nishisato, S. Decision latencies of same and different judgments. Perception and Psychophysics, 1968, 3, 121-130.

Biondi, E., & Biondi, L. The sampling of speech sounds as a new means of making speech intelligible to the deaf. Alta Frequenza, 1968, 37, 728-739.

Bixler, R., Foulke, E., Amster, C., & Nolan, C. The comprehension of rapid speech by the blind: Part I. (Cooperative Research Project No. 1005) Washington, D.C.: United States Department of Health, Education, and Welfare, Office of Education, 1961.

Black, J. W. Practical limits of speeded speech. United States Air Force Technical Report No. 5917. Dayton: Wright-Patterson Air Force Base, 1950. (Cited by Garvey, W. D., & Henneman, R. H.)

Black, J. W. Aural perception of sentences of different lengths. Quarterly Journal of Speech, 1961, 47, 51-53.

Bloom, B. J. (Ed.) Taxonomy of educational objectives, the classification of educational goals handbook I: Cognitive domain. New York: David McKay, 1964.

Bloomfield, L., & Barnhart, C. L. Let's read: a linguistic approach. Detroit: Wayne State University Press, 1961.

Bocca, E. Binaural hearing: Another approach. Laryngol, 1955, 65, 1164-1171.

Bocca, E. Factors influencing binaural integration of periodically switched messages. Acta Oto-Laryngologica, 1961, 53, 142-144.

Bocca, E. Distorted speech tests. In A. B. Graham (Ed.), Sensorineural hearing processes and disorders. Boston: Little, Brown, 1967.

Bocca, E., & Calearo, C. Central hearing processes. In J. Jerger (Ed.), Modern developments in audiology. New York: Academic Press, 1963.

Bocca, E., Calearo, C., Cassinari, V., & Magliavacca, F. Testing cortical hearing in temporal lobe tumors. Acta Oto-Laryngologica, 1955, 45, 289-303.

Boesch, E., Baltes, P., & Schmidth, L. Preferences for different auditory stimulus sequences in various age groups. Psychonomic Science, 1968, 10, 205-206.

Bogart, B. The vobanc. A two-to-one speech bandwidth reduction system. Journal of the Acoustical Society of America, 1956, 28, 399-404.

Boomer, D., & Dittman, A. Hesitation pauses and juncture pauses in speech. Language and Speech, 1962, 5, 215-226.

Boullion, J-P. Resultats de reeducation auditires effecfue es par des professeurs de l'Institut de Jeunes Sourd de Paris. Proceedings of Colloque Internationale, Paris, 1961.

Brent, S. B. Linguistic unity, list length, and rate of presentation in serial anticipating learning. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 70-79.

Bricker, P., & Pruzansky, S. Effects of stimulus content and duration on talker identification. Journal of the Acoustical Society of America, 1966, 40, 1441-1449.

Broadbent, D. Growing points in multi-channel communication. Journal of the Acoustical Society of America, 1956, 28, 533-535.

Broadbent, D. Perception and communication. New York: Pergamon Press, 1958.

Brown, J. I., & Carlsen, G. R. Brown-Carlsen listening comprehension test: Evaluation and adjustment series. New York: Harcourt, Brace & World, 1955.

Brown, S. W., Guilford, J. P., & Hoepfner, R. A factor analysis of semantic memory abilities. (Report No. 37) Psychological Laboratory, University of Southern California, Los Angeles, 1966.

Bruner, J. S. Personality dynamics and the process of perceiving. In R. R. Blake & G. V. Ramsay (Eds.), Perception, An approach to personality. New York: Ronald Press, 1951.

Bryan, M., & Tempest, W. Precision audiometry. Acta Oto-Laryngologica, 1967, 64, 205-212.

Buehr, W. Bread the staff of life. New York: William Morrow, 1959.

Calearo, C. Binaural summation in lesions of the temporal lobe. Acta Oto-Laryngologica, 1957, 47, 392-397.

Calearo, C. Cortical hearing tests and cerebral dominance. Acta Oto-Laryngologica, 1963, 56, 17-26.

Calearo, C., & di Mitri, T. Sulla intelligibilita della voce periodicemente alternata da un orchio all altro. Congress di Audiologica, Padova, Italy, 1958.

Calearo, C., & Lazzaroni, A. Speech intelligibility in relation to the speed of the message. Laryngoscope, 1957, 67, 410-419.

Calearo, C., Teatini, G. P., & Pestalozza, G. Speech intelligibility in the presence of interrupted noise. Journal of Auditory Research, 1962, 2, 179-187.

Calhoun, S. Influence of syllabic length and rate of auditory presentation on ability to reproduce disconnected word lists. Journal of Experimental Psychology, 1935, 18, 612-620.

Campanella, S. J. Signal analysis of speech time compression techniques. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967.

Carey, P. W. Delayed auditory feedback and the shadowing response. Unpublished doctoral dissertation, Harvard University, 1968.

Carhart, R. Future horizons in audiologic diagnosis. Annals of Otology, Rhinology, and Laryngology, 1968, 77, 706-716.

Carhart, R. Special hearing tests for otoneurologic diagnosis. Archives of Otolaryngology, 1969, 89, 64-66.

Carlson, J. Digitalized phase vocoder. AFCRL and Institute of Electrical and Electronic Engineers Audio and Electro-acoustics 1967 Conference on Speech Communication and Processing. Bedford: OAR, United States Air Force, 1967.

Carraway, B., & Carhart, R. Influence of compressor action on speech intelligibility. Journal of the Acoustical Society of America, 1967, 41, 1424-1433.

Carrell, J., & Tiffany, W. Phonetics: Theory and application to speech improvement. New York: McGraw-Hill, 1960.

Carroll, J. B. Language and thought. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Carroll, J. B. Problems of measuring speech rate. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967.

Carroll, J. B., & Cramer, H. L. The intelligibility of time-compressed speech. (Final Report Contract No. OE-7-31-0370-271) Washington, D. C.: United States Department of Health, Education, and Welfare, Office of Education, 1968.

Cavanaugh, F., & Crandall, E. Meet the presidents. Philadelphia: Macrae Smith, 1962.

Center for Rate-Controlled Recordings Newsletter. Louisville: University of Louisville.

Chang, S. Two schemes of speech compression system. *Journal of the Acoustical Society of America*, 1956, *28*, 565-572.

Chapman, W. Speech compression by tape loop and by computer. In, *Proceedings of the Louisville conference on time-compressed speech*. Louisville: University of Louisville, 1967.

Chatterjea, R. Time gap in the estimation of short duration with auditory presentation. *Australian Journal of Psychology*, 1959, *11*, 182-190.

Chomsky, N., & Halle, M. *Sound pattern of English*. New York: Harper & Row, 1968.

Central institute for the deaf manual for auditory tests W-1 and W-2. St. Louis: Central Institute for the Deaf, 1957.

Clark, C. R. (Ed.) *Negro heritage series*. Nashville, Tenn.: George Peabody College, Institute on Mental Retardation and Intellectual Development, 1968.

Clark, C. R., & Woodcock, R. W. Standardized listening passages. *Institute on Mental Retardation and Intellectual Development Papers and Reports*, *4*(2), 1967.

Cody, D., Griffing, T., & Taylor, W. Assessment of the newer tests of auditory function. *Archives of Otolaryngology*, 1968, *77*, 686-705.

Comer, D. J. The use of waveform asymmetry to identify voiced sounds. *Institute of Electrical and Electronic Engineers Transactions on Audio and Electro-acoustics*, 1966. (Pp. 500-506)

Compton, A. Effects of filtering and vocal duration upon the identification of speakers aurally. *Journal of the Acoustical Society of America*, 1964, *35*, 1748-1753.

Conant, R., & Ashby, W. Information, transmission, and redundancy in regulatory systems. *Biological Computer Laboratory Report No. 67.2*. Champaign, Ill.: University of Illinois, Biological Computer Laboratory, 1968.

Condon, E. F. An analysis of the differences between good and poor listeners in grades nine, eleven, and thirteen. (Unpublished doctoral dissertation, University of Kansas, 1965) *Dissertation Abstracts*, 1965, *26*, 3106.

Cooke, J., & Beard, S. Speech intelligibility for space vehicles using nitrogen or helium as the inert gas. *Journal of the Acoustical Society of America*, 1966, *40*, 1450-1453.

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., & Gerstman, L. J. Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America*, 1952, *24*(6), 597-606.

Cooper, F. S., Liberman, A. M., & Borst, J. M. The interconversion of audible and visible patterns as a basis for research in the perception of speech. *Proceedings of the National Academy of Sciences*, 1951, *37*(5), 318-325.

Corliss, E., Burnett, E., Kobal, M., & Bassin, M. The relative importance of frequency distortion and changes in time constants in the intelligibility of speech. *Institute of Electrical and Electronic Engineers Transactions on Audio and Electro-acoustics*, 1968, *AU-15*, 36-39.

Cornett, R. Cued speech. *American Annals of the Deaf*, 1967, *112*, 3-13.

Cramer, H. L. Intelligibility of compressed speech as a function of degree and direction of delay of presentation from one ear to the other. Paper presented at the American Psychological Association, Chicago, September 1965.

Cramer, H. L. The intelligibility of time-compressed speech. In, *Proceedings of the Louisville conference on time-compressed speech*. Louisville: University of Louisville, 1967.

Cramer, H. L. The intelligibility of time-compressed speech. Unpublished doctoral dissertation, Harvard University, 1968.

Cramer, H. L. An introduction to speech compression techniques. In, *Proceedings of the second Louisville conference on rate and/or frequency-controlled speech*. Louisville: University of Louisville, Perceptual Alternatives Laboratory, 1970.

Cramer, H. L., & Talambiras, R. The braided speech method of time compressing speech. In, *Proceedings of the second Louisville conference on rate and/or frequency-controlled speech*. Louisville: University of Louisville, Perceptual Alternatives Laboratory, 1970.

Cronbach, L. J. Essentials of psychological testing. (2nd ed.) New York: Harper, 1960.

Cross, P. K. The junior college student: A research description. Princeton, N. J.: Educational Testing Service, 1968.

Dale, E., & Chall, J. S. A formula for predicting readability. Educational Research Bulletin, 1948, 27, 11-20.

Daniloff, R., Shriner, T., & Zemlin, W. Intelligibility of vowels altered in duration and frequency. Journal of the Acoustical Society of America, 1968, 44, 700-707. (a)

Daniloff, R., Shriner, T., & Zemlin, W. The intelligibility of vowels altered in time and frequency. Paper presented at the 75th Semi-annual Meeting of the Acoustical Society of America, Ottawa, May 1968. (b)

Daniloff, R., Shriner, T., & Zemlin, W. A study of consonantal confusions with time and frequency distorted speech signals. Champaign, Ill.: University of Illinois, 1968. (c)

Darley, F. L. A normative study of oral reading rate. Unpublished master's thesis, State University of Iowa, 1940.

David, E. Naturalness and distortion in speech processing devices. Journal of the Acoustical Society of America, 1956, 28, 586-589.

Davis, H. (Ed.), & Silverman, S. R. (Co-ed.), Hearing and deafness. (Rev. ed.) New York: Holt, Rinehart & Winston, 1964.

deHoop, W. Listening comprehension of cerebral palsied and other crippled children as a function of two speaking rates. Exceptional Children, 1965, 31, 232-240.

deHoop, W. Effects and interaction effects of speaking rate, visual limitation, and intelligence level on aural acquisition and retention of sentences. (Unpublished doctoral dissertation, University of Georgia, 1965) Dissertation Abstracts, 1966, 26, 3752-3753.

deHoop, W. Effects and interaction effects of speaking role; visual limitation, and intelligence level on aural acquisition and retention of sentences. In, Proceedings of the Louisville conference on time-compressed speech. Louisville, University of Louisville, 1967.

Delattre, P., Cooper, F. S., Liberman, A. M., & Gerstman, L. J. Speech synthesis as a research technique. In, Proceedings of the VIIth international congress of linguists. London: Titus Wilson, 1956.

Delattre, P., Liberman, A. M., & Cooper, F. S. Acoustic loci and transitional cues for consonants. Journal of the Acoustical Society of America, 1955, 27, 769-773.

Denes, P. Effect of duration on the perception of voicing. Journal of the Acoustical Society of America, 1955, 27, 761-764.

Denes, P. An introduction to problems of sensory aids for the deaf. Institute of Electrical and Electronic Engineers Transactions on Audio and Electro-acoustics, 1969.

deQuiros, J. B. Accelerated speech audiometry, an examination of test results. Translations of the Beltone Institute for Hearing Research, No. 17, 1964.

Derleth, A. A battle over the teacups. In R. C. Pooley, I. C. Poley, J. C. Leyda, & L. J. Zellhoefer (Eds.), Good times through literature. Chicago: Scott, Foresman, 1957. (Pp. 40-43)

DeVito, J. A. CLOZE procedure. Today's Speech, 1967, 15(1).

Diagnostic reading tests, section II, comprehension: Silent and auditory. Mountain Home, N. C.: Committee on Diagnostic Reading, 1963.

Diamond, R., & Kinney, R. The place of speech compression in academic study. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967.

Diehl, C. F., White, R. C., & Burk, K. Rate and communication. Speech Monographs, 1959, 26, 229-232.

Discussion of residual hearing and auditory recording. American Annals of the Deaf, 1968, 113, 307-310.

Discussion of the speech code and visual speech. American Annals of the Deaf, 1967, 113, 203-204.

Dudley, H. Signaling system. Washington, D. C.: United States Patent Office, May 3, 1938. U. S. Patent No. 2,115,803.

Dudley, H. The vocoder. *Bell Laboratories Record*, 1939, *18*, 122-126.

Duffy, E. *Activation and behavior*. New York: John Wiley & Sons, 1962.

Duker, S. *Listening bibliography*. New York: Scarecrow Press, 1964.

Duker, S. Listening and reading. *Elementary School Journal*, 1965, *65*, 321-329.

Dunn, L. M. *Peabody picture vocabulary test*. Circle Pines, Minn.: American Guidance Service, 1959.

Durost, W. N., Bixler, H. H., Hildreth, G. H., Lund, K. W., & Wrightstone, J. W. *Metropolitan achievement tests*. New York: Harcourt, Brace & World, 1962.

Egan, J. P. Articulation testing methods. *Laryngoscope*, 1948, *58*, 955-991.

Electronics research directorate. In, *Proceedings of the seminar on speech compression and processing*. Bedford: United States Air Force Cambridge Research Center, 1959.

Ellis, N. R. (Ed.) *Handbook of mental deficiency*. New York: McGraw-Hill, 1963.

Enc, M. E. The effect of two different wpm listening rates on learning and retention of blind school children. (Unpublished doctoral dissertation, University of Illinois, 1958) *Dissertation Abstracts*, 1959, *19*, 2512.

Enc, M. E., & Stolurow, L. M. A comparison of the effects of two recording speeds on learning and retention. *New Outlook for the Blind*, 1960, *54*, 39-48.

Endres, W. Experiments concerning the comparison of some different methods of speech bandwidth compression using only parts of the time-frequency spectrum of speech. *AFCRL and Institute of Electrical and Electronic Engineers Audio and Electro-acoustics Conference on Speech Communication and Processing*. Bedford: OAR, United States Air Force, 1967.

Endres, W. On the experimental evaluation of the interior redundancy in speech. (No. B-3-10) *Reports of the 6th International Congress on Acoustics*, 1968.

Ernest, C. Listening comprehension as a function of type of material and rate of presentation. *Speech Monographs*, 1968, *35*, 154-158.

Ethington, D. The readability of braille as a function of three spacing variables. Unpublished master's thesis, University of Kentucky, 1956.

Fairbanks, G. Experimental studies of time compression of speech. *Journal of the Acoustical Society of America*, 1956, *28*.

Fairbanks, G. Test of phonemic differentiation: The rhyme test. *Journal of the Acoustical Society of America*, 1958, *30*, 596-599.

Fairbanks, G. *Voice and articulation drillbook*. New York: Harper & Brothers, 1960.

Fairbanks, G., Everitt, W. L., & Jaeger, R. P. Method for time or frequency compression-expansion of speech. *Transactions of the Institute of Radio Engineers, Professional Group on Audio*, 1954, *AU-2*, 7-12. Also in: *Convention Record of the Institute of Radio Engineers, 1953 National Convention Part 8: Information Theory*, 1953. (Pp. 120-124)

Fairbanks, G., Everitt, W. L., & Jaeger, R. P. *Recording device*. Washington, D. C.: United States Patent Office, May 12, 1959. U. S. Patent No. 2,886,650.

Fairbanks, G., Guttman, N., & Miron, M. S. Auditory comprehension in relation to listening rate and selective verbal redundancy. *Journal of Speech and Hearing Disorders*, 1957, *22*, 23-32. (a)

Fairbanks, G., Guttman, N., & Miron, M. S. Auditory comprehension of repeated high-speed messages. *Journal of Speech and Hearing Disorders*, 1957, *22*, 20-22. (b)

Fairbanks, G., Guttman, N., & Miron, M. S. Effects of time compression upon the comprehension of connected speech. *Journal of Speech and Hearing Disorders*, 1957, *22*, 10-19. (c)

Fairbanks, G., & Kodman, F., Jr. Word intelligibility as a function of time compression. *Journal of the Acoustical Society of America*, 1957, *29*, 636-641.

Fant, G., & Risberg, A. Evaluation of speech compression systems. *Speech Trans. Laboratory Quarterly Progress Status Report*, 1957, *12*, 636-641.

Faroqui, M., & Parameswaran, E. Effect of interval between signals on temporal judgment. Canadian Journal of Psychology, 1967, 20, 479-495.

Fawcett, A. E. The effect of training in listening upon the listening skills of intermediate grade children. (Unpublished doctoral dissertation, University of Pittsburgh, 1963) Dissertation Abstracts, 1965, 25, 7108-7109.

Feldman, R., & Reger, S. Relations among hearing, reaction time, and age. Journal of Speech and Hearing Research, 1967, 10, 479-495.

Fergen, G. K. Listening comprehension at controlled rates for children in Grades IV, V, and VI. (Unpublished doctoral dissertation, University of Missouri, 1954) Dissertation Abstracts, 1955, 15, 89.

Ferguson, G. A. On learning and human ability. Canadian Journal of Psychology, 1956, 10, 121-131.

Ferguson, G. A. Nonparametric trend analysis. Montreal: McGill University Press, 1965.

Ferguson, G. A. Statistical analysis in psychology and education. New York: McGraw-Hill, 1966. (Pp. 317-319)

Flanagan, J. Speech analysis synthesis and perception. New York: Academic Press, 1965. (Chs. 7-8)

Flanagan, J., & Golden, R. Phase vocoder. Bell System Technical Journal, 1966, 1493-1509.

Fleishman, E. Factor structure in relation to task difficulty in psychomotor performance. Educational and Psychological Measurement, 1957, 17, 522-532.

Flesch, R. A new readability yardstick. Journal of Applied Psychology, 1948, 32, 231-233.

Flesch, R. How to test readability. New York: Harper, 1951.

Fletcher, H. Speech and hearing. New York: Van Nostrand Co., 1929. (Pp. 293-294)

Fletcher, H. Speech and hearing in communications. New York: Van Nostrand Co., 1965. (Chs. 15-19)

Flowers, A., & Costello, M. Responses to distorted speech of children with severe articulation disorders. Journal of Auditory Research, 1963, 3, 133-140.

Foder, J. A., & Bever, T. G. The psychological reality of linguistic segments. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 414-420.

Foulke, E. A comparison of two methods of compressing speech. Paper presented at the Southeastern Psychological Association, Louisville, March 1962.

Foulke, E. The comprehension of rapid speech by the blind: Part II. (Cooperative Research Project No. 1370) Washington, D. C.: United States Department of Health, Education, and Welfare, Office of Education, 1964. (a)

Foulke, E. Transfer of a complex perceptual skill. Perceptual and Motor Skills, 1964, 18, 733-740. (b)

Foulke, E. Comparison of comprehension of two forms of compressed speech. Journal of Exceptional Children, 1966, 33, 169-173. (a)

Foulke, E. The retention of information presented at an accelerated word rate. International Journal for the Education of the Blind, 1966, 14(1), 11-15. (b)

Foulke, E. A survey of the acceptability of rapid speech by the blind. The New Outlook for the Blind, 1966, 60, 261-265. (c)

Foulke, E. The influence of a reader's voice and style of reading on comprehension of time-compressed speech. The New Outlook for the Blind, 1967, 61(3), 65-68. (a)

Foulke, E. (Ed.) Proceedings of the Louisville Conference on Time-Compressed Speech. Louisville: University of Louisville, 1967. (b)

Foulke, E. Time-compressed recorded speech and faster aural reading. Blindness 1967, American Association of Workers for the Blind Annual, 1967, 11-19. (c)

Foulke, E. Listening comprehension as a function of word rate. Journal of Communication, 1968, 18, 198-206. (a)

Foulke, E. The perception of time-compressed speech. Paper presented at the meeting of the Perception of Language Conference, Pittsburgh, January 1968. (b)

Foulke, E. The comprehension of rapid speech by the blind: Part III. (Final Progress Report, Cooperative Research Project No. 2430) Washington, D. C.: United States Department of Health, Education, and Welfare, Office of Education, 1969.

Foulke, E., Amster, C. H., Nolan, C. Y., & Bixler, R. H. The comprehension of rapid speech by the blind. Exceptional Children, 1962, 29, 134-141.

Foulke, E., & Bixler, R. H. Speech compression. The New Beacon, 1963, 47, 241-242.

Foulke, E., & Sticht, T. Listening rate preferences of college students for literary material of moderate difficulty. Journal of Auditory Research, 1966, 6, 397-401.

Foulke, E., & Sticht, T. The intelligibility and comprehension of time-compressed speech. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967. (a)

Foulke, E., & Sticht, T. A review of research on time-compressed speech. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967. (b)

Foulke, E., & Sticht, T. Review of research on the intelligibility and comprehension of accelerated speech. Psychological Bulletin, 1969, 72(1), 50-62.

Franke, P. A preliminary study validating the measurement of speech rate variation. Unpublished master's thesis, University of Iowa, 1939.

Fraser, D. Decay of immediate memory with age. Nature, 1958, 182, 1163.

French, N. R., & Steinberg, J. C. Factors governing the intelligibility of speech sounds. Journal of the Acoustical Society of America, 1947, 19(1), 90-119.

French, N. R., & Zinn, M. K. Method of and apparatus for reducing width of transmission bands. Washington, D. C.: United States Patent Office, May 29, 1928. U. S. Patent No. 1,671,151.

Freund, B. Method of and apparatus for varying the length of sound records. Washington, D. C.: United States Patent Office, April 9, 1935. U. S. Patent No. 1,996,958.

Friedman, H. L. Overview of compressed speech research, 1963-1967. Silver Spring, Md.: American Institutes for Research, 1967.

Friedman, H. L., Graae, C. N., & Orr, D. B. Further research on speeded speech as an educational medium--effects of listening aids and self-pacing on comprehension and the use of compressed speech for review. Report No. R67-2. Silver Spring, Md.: American Institutes for Research, 1967.

Friedman, H. L., & Johnson, R. L. Compressed speech: correlates of listening ability. Journal of Communication, 1968, 18, 207-218.

Friedman, H. L., & Johnson, R. L. Time-compressed speech as an educational medium: Studies of stimulus characteristics and individual differences. Final Report No. R69-14. Silver Spring, Md.: American Institutes for Research, 1969.

Friedman, H. L., & Orr, D. B. Recent research in the training of compressed speech comprehension. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967.

Friedman, H. L., Orr, D. B., Freedle, R. O., & Norris, C. M. Further research on speeded speech as an educational medium. Progress Report No. 2. Silver Spring, Md.: American Institutes for Research, 1966.

Friedman, H. L., Orr, D. B., & Norris, C. M. Further research on speeded speech as an educational medium: The use of listening aids. Progress Report No. 3. Silver Spring, Md.: American Institutes for Research, 1966.

Fries, C. C. The structure of English. New York: Harcourt, Brace, 1952.

Fry, D. Duration and intensity as physical correlates of linguistic stress. Journal of the Acoustical Society of America, 1955, 27, 765-768.

Fujimura, O. An approximation to voice aperiodicity. AFCRL and Institute of Electrical and Electronic Engineers Audio and Electroacoustics 1967 Conference on Speech Communication and Processing. Bedford: OAR, United States Air Force, 1967.

Fussel, G., & Thomas, C. Speech compression by the removal of periodic redundancy. AFCRL and Institute of Electrical and Electronic Engineers Audio and Electro-acoustics 1967 Conference on Speech Communication and Processing. Bedford: OAR, United States Air Force, 1967.

Gabor, D. Theory of communication. Journal of the Institute of Electrical Engineers, 1946, 93(3), 429-457.

Gabor, D. New possibilities in speech transmission. Journal of the Institute of Electrical Engineers, 1947, 94(3), 369-387.

Gabor, D. Improvements in and relating to transmission, recording, and reproduction of sound. London: The Patent Office, January 18, 1949. British Patent No. 616,156.

Gabor, D. Improvements in and relating to the transmission of speech. London: The Patent Office, November 1, 1950. British Patent No. 645,542.

Gabor, D. Bandwidth compression and expansion system for transmission of speech. A description of a speech compressor. Washington, D. C.: United States Patent Office, May 11, 1965. U. S. Patent No. 3,183,310.

Gabrilovitch, L. Method and device for reducing and multiplying acoustic frequencies. London: The Patent Office, September 12, 1938. British Patent No. 491,937.

Gabrilovitch, L. Method and device for reducing and increasing audio frequencies. Washington, D. C.: United States Patent Office, August 22, 1939.

Garvey, W. D. Duration factors in speech intelligibility. Unpublished master's thesis, University of Virginia, 1949.

Garvey, W. D. An experimental investigation of the intelligibility of speeded speech. Unpublished doctoral dissertation, University of Virginia, 1951.

Garvey, W. D. The intelligibility of abbreviated speech patterns. Quarterly Journal of Speech, 1953, 39, 296-306. (a)

Garvey, W. D. The intelligibility of speeded speech. Journal of Experimental Psychology, 1953, 45, 102-108. (b)

Garvey, W. D., & Henneman, R. H. Practical limits of speeded speech. United States Air Force Technical Report No. 5917. Dayton: Wright-Patterson Air Force Base, 1950.

Gates, A. I. Gates reading survey. New York: Columbia University Teachers College, Bureau of Publications.

Gazdag, J. A method of decoding speech. Technical Report No. 9, 1966, University of Illinois, Contract AF 7-66, Electrical Engineering Research Laboratory, Engineering Experimental Station.

Gemelli, A. The minimum duration of a phoneme sufficient for its perception. In A. Gemmelli & G. Pastori (Eds.), L'analisi elettroacustica del linguaggio. Milan, Italy: Societa Editrice "Vita e Pensiero," 1934. (Pp. 149-162)

George, R. Difficulty of material and retention of compressed speech. Unpublished doctoral dissertation, Indiana State University, 1969.

Gerber, S. E. Dichotic and diotic presentation of speeded speech. Journal of Communication, 1968, 18, 272-282.

Gerber, S. E. Monotic versus diotic presentation of dichotic speeded speech. Journal of Communication, 1969, 19, 325-332.

Gerber, S. E. Scientific computation in speech research. Computer Studies in Humanities and Verbal Behavior, in press.

Gerber, S. E., & Scott, R. J. Dichotic speech-time compression. In, Proceedings of the second Louisville conference on rate and/or frequency-controlled speech. Louisville: University of Louisville, Perceptual Alternatives Laboratory, 1970.

Gibbons, E. W., Winchester, R. A., & Krebs, D. F. The variability of oral reading rate. Journal of Speech and Hearing Disorders, 1958, 23, 591-593.

Gilber, J. H. V. Rate alterations in oral reading. Unpublished master's thesis, Purdue University, 1965.

Gilmour, W. D. A general purpose phonemic transcriptor. Institute for Electrical Engineers N. P. L. Conference on Pattern Recognition, 1968, 154-167.

Gold, B. Techniques for speech bandwidth compression, using combinations of channel vocoder and formant vocoder. Journal of the Acoustical Society of America, 1952, 38, 2-10.

Gold, B., & Radar, C. The channel vocoder. *Institute of Electrical and Electronic Engineers Transaction on Audio and Electro-acoustics*, 1968, *AU-15*, 148-160.

Golden, R. M. Digital computer simulation of a sampled-data voice-excited vocoder. *Journal of the Acoustical Society of America*, 1963, *35*, 1358-1367.

Golden, R. M. Improving naturalness and intelligibility, using vocoder techniques. *Journal of the Acoustical Society of America*, 1966, *40*, 621-624.

Golden, R. M. Comments on "technique for correcting helium speech distortion"/ W. Stover. *Journal of the Acoustical Society of* America, 1967, *41*/ *42*, 70-74/ 514.

Golden, R. M. Vocoder filter design: Practical considerations. *Journal of the Acoustical Society of America*, 1968, *43*, 803-810.

Goldhaber, G. Listener comprehension of compressed speech when the difficulty level of the content and the sex of the listener are varied. Unpublished master's thesis, University of Maryland, 1967.

Goldhaber, G., & Weaver, C. Listener comprehension of compressed speech when the difficulty, rate of presentation, and sex of the listener are varied. *Speech Monographs*, 1968, *35*, 20-25.

Goldman, R. M. Time compression and expansion of recorded speech. Paper presented at the meeting of the Acoustical Society of America, Los Angeles, November 1966.

Goldman-Eisler, F. The determinants of the rate of speech output and their mutual relations. *Journal of Psychosomatic Research*, 1956, *1*, 137-143.

Goldman-Eisler, F. Speech production and the predictability of works in context. *Quarterly Journal of Experimental Psychology*, 1958, *10*(2), 96-106.

Goldman-Eisler, F. A comparative study of two hesitation phenomena. *Language and Speech*, 1961, *4*, 18-26. (a)

Goldman-Eisler, F. Continuity of speech utterance, its determinants and its significance. *Language and Speech*, 1961, *4*, 220-231. (b)

Goldman-Eisler, F. The distribution of pause durations in speech. *Language and Speech*, 1961, *4*, 232-237. (c)

Goldman-Eisler, F. Psycholinguistics: Experiments in spontaneous speech. New York: Academic Press, 1968.

Goldstein, H. Reading and listening comprehension at various controlled rates. Teachers College Contributions to Education No. 821. New York: Bureau of Publications, Columbia University, Teachers College, 1940.

Goodman-Malmuth, L. An experimental study of the effects of rate of speaking upon listenability. Unpublished doctoral dissertation, University of California, 1956. Also in: Speech Monographs, 1957, 24, 89-90.

Graham, W. Time factor alteration of recorded sound while maintaining frequency constants. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967.

Graham, W. The Graham compressor, a technical development of the Fairbanks method. In, Proceedings of the second Louisville conference on rate and/or frequency-controlled speech. Louisville: University of Louisville, Perceptual Alternatives Laboratory, 1970.

Grandine, J. D., & Hagan, T. G. A parallel/sequential, stored-program hybrid signal processor. Simulation, January, 1965.

Grimm, W. Perception of segments of English-spoken consonant-vowel syllables. Journal of the Acoustical Society of America, 1966, 40, 1454-1462.

Gropper, R. Comprehension of narrative passages by fourth-grade children as a function of listening rate and eleven prediction variables. Unpublished manuscript, George Peabody College, 1969.

Guilford, J. P. Fundamental statistics in psychology and education. (3rd ed.) New York: McGraw-Hill, 1956.

Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Guilford, J. P. Intelligence has three facets. Science, 1968, 160, 615-620.

Guttman, N., & Nelson, J. An instrument that creates some artificial speech spectra for the severely hard of hearing. American Annals of the Deaf, 1968, 113, 295-302.

Guttman, N., & van Bergeijk, W. Frequency division in real time as a means of counteracting high frequency hearing loss. In, Proceedings of the 4th international congress of the societe internationale d'audiologie, Padua, Italy, 1958.

Hansen, R. An experimental investigation of the effects of different rates of speaking intelligibility. Unpublished master's thesis, Ohio State University, 1950.

Harley, R. An experimental program in compressed speech at the Tennessee School for the Blind. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967.

Harris, A. J. How to increase reading ability. New York: Longman Green, 1947.

Harris, J. Combinations of distorted speech: The twenty-five percent safety factor by multiple-cueing. Archives of Otolaryngology, 1960, 70, 131-145.

Harris, J., Haines, H., & Myers, C. The importance of hearing at 3 KC for understanding speeded speech. Laryngoscope, 1960, 70, 131-145.

Hartlage, L. Differences in listening comprehension between blind and sighted subjects. International Journal for the Education of the Blind, 1963, 13, 1-6.

Harwood, K. A. Listenability and rate of presentation. Speech Monographs, 1955, 22, 57-59.

Haviland, R. Space broadcasting--how, when, and why. Paper presented at a meeting of the United Nations Educational, Scientific, and Cultural Organization Conference on Explorations and Peaceful Uses of Outer Space, Vienna, 1968.

Hecker, M., Stevens, K., & Williams, C. Measurements of reaction time in intelligibility tests. Journal of the Acoustical Society of America, 1966, 39, 1188-1189.

Hecker, M., von Bismark, G., & Williams, C. Automatic evaluation of time-varying communications systems. Institute of Electrical and Electronic Engineers Transactions on Audio and Electro-acoustics, 1968, AU-15, 100-106.

Hedlund, J. L. Army classification scores: their meaning and clinical implication. Paper for the Conference of Army Clinical Psychologists, Department of the Army, Washington, D. C., September 1969.

Henneman, R. A technique for the investigation of the practical limits of speeded speech. American Psychologist, 1949, 4, 304.

Henry, W. G., Jr. Recognition of time-compressed speech as a function of word length and frequency of usage. (Unpublished doctoral dissertation, Indiana University, 1966) Dissertation Abstracts, 1967, 27, 3310A-3311A.

Henze, R. Experimentelle untersuchunger zur phanomenologie der sprechgeschwindigkeit. Zeitschrift fur Psychologie, 1953, 1, 214-243.

Hill, M. The speech intelligibility in presbycusis under conditions of simultaneous time and frequency distortion. Unpublished doctoral dissertation, Wayne State University, 1964.

Hirsch, I. Auditory perception of temporal order. Journal of the Acoustical Society of America, 1959, 31, 759-767.

Hirsch, I. Audition in relation to perception of speech. Brain Function. Vol. III: Speech, Language, and Communication. Berkley: University of California Press, 1966.

Hirsch, I. Information processing in input channels for speech and language: The significance of serial order of stimuli. In C. H. Millikan & F. L. Darley (Eds.), Brain mechanisms underlying speech and language. New York: Grune & Stratton, 1967.

Hoepfner, R., Nihira, K., & Guilford, J. P. Intellectual abilities of symbolic and semantic judgment. Psychological Monographs, 1966, 80, 1-47.

Hogan, D. L., & Scott, R. J. A hybrid PDP-1 system for speech research. Proceedings of the 1963 Annual Meeting of the Digital Equipment Users Society, Livermore, California, 1963, 183-191.

Hogben, M. An investigation of compression in compressed speech. Unpublished master's thesis, University of Utah, 1967.

Hollywell, N., & Harvey. Helium speech. Journal of the Acoustical Society of America, 1964, 36, 210-211.

Hopkins, F. W. An experimental study of the effects of repetitive compressed speech on listening comprehension. Unpublished doctoral dissertation, University of Maryland, 1969.

Hopkinson, N. T. Combined effects of interruption and interaural alternation on speech intelligibility. Language and Speech, 1964, 10, 210-211.

Hudgins, C. F., Hawkins, J. E., Karlin, J. E., & Stevens, S. S. The development of recorded auditory tests for measuring hearing loss for speech. Laryngoscope, 1947, 57, 57-59.

Huggins, A. Distortion of the temporal pattern of speech: Interruption and alternation. Journal of the Acoustical Society of America, 1965, 36, 1055-1064.

Hunter, E. Problems of diver communication. AFCRL and Institute of Electrical and Electronic Engineers Audio and Electro-acoustics 1967 Conference on Speech Communication and Processing. Bedford: OAR, United States Air Force, 1967.

Hutton, C. L., Jr. A psychophysical study of speech rate. (Unpublished doctoral dissertation, University of Illinois, 1954) Dissertation Abstracts, 1955, 15, 168.

Indiresan, P. Interrupted speech and the possibility of increasing communication effectiveness. Journal of the Acoustical Society of America, 1963, 35, 405-408.

Iverson, L. Time compression. International Journal for the Education of the Blind, 1958, 5(4), 78-79.

Jahnke, J. Presentation rate and the serial position effect of immediate serial recall. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 608-612.

Jamison, D., Ferraz, J., & Torquato, J. Alternatives for instructional broadcast satellites. Institute of Electrical Engineers Transactions on Broadcasting, 1969, BC-15, 1-6.

Jamison, D., Jamison, M., & Hewlett, S. Satellite radio--better than ETV. Astronautics and Aeronautics, 1969, 7, 92-97.

Jankowski, W., & Iwankiewicz, S. Some problems of functional studies on the efficiency of central hearing pathways. Paper presented at a meeting of the World Congress of the Deaf, Stockholm, Sweden, 1963.

Jerger, J. Observations on auditory lesions in the central auditory pathways. Archives of Otolaryngology, 1960, 71, 797-806.

Jerger, J. Auditory tests for disorders of the central auditory mechanism. In W. S. Fields & B. R. Alford (Eds.), Neurological aspects of auditory and vestibular disorders. Springfield: C. C. Thomas, 1964.

Jerger, J., & Jerger, S. Psychoacoustic comparison of chochlear and VIIIth nerve disorders. Journal of Speech and Hearing Research, 1967, 10, 659-688.

Jerger, J., Mier, M., Boshes, B., & Canter, G. Auditory behavior in parkinsonism. Acta Oto-Laryngologica, 1960, 52, 541-550.

Jester, R. E. Comprehension of connected discourse as a function of individual differences and rate and modality of presentation. (Unpublished doctoral dissertation, University of Utah, 1966) Dissertation Abstracts, 1966, 27, 957.

Jester, R. E., & Travers, R. M. W. Comprehension as a function of rate and modality of presentation. Paper presented at the meeting of the American Psychological Association, Chicago September 1965.

Jester, R. E., & Travers, R. M. W. Comprehension of connected meaningful discourse as a function of rate and mode of presentation. Journal of Educational Research, 1966, 59(7), 297-302.

Jester, R. E., & Travers, R. M. W. The effect of various presentation patterns on the comprehension of speeded speech. American Educational Research Journal, 1967, 4, 353-360.

Johansson, B. A new coding amplifier system for the hard of hearing. Proceedings of the 3rd International Congress on Acoustics, 1961, 2, 655-657.

Johansson, B. The use of the transposer for the management of the deaf child. International Audiology, 1966, 5, 362-372.

Johannson, B., & Wedenberg, E. Speech compression techniques applied to aids for the hard of hearing. (1) Hearing aids. Speech Trans. Laboratory Quarterly Progress Status Report, 1960, 1, 25-26.

Johnson, R. L., Friedman, H. L., & Stuart, C. Effect of temporal spacing on the recall of time-compressed sentences. (1970, in preparation)

Johnson, W. Measurements of oral reading and speaking rate and disfluency of adult male and female stutterers and nonstutterers. Journal of Speech and Hearing Disorders, Monograph Supplement, 1961, 7, 1-20.

Johnson, W., Darley, F., & Spriestersbach, D. C. Diagnostic methods in speech pathology. New York: Harper & Row, 1963. (Pp. 202-203)

Karlin, J. E., et al. Auditory tests of the ability to hear speech in noise. OSRD Report No. 3516, Psycho-acoustic Laboratory, Harvard University, 1944.

Kato, Y., Fujimura, O., & Maeda, S. A vocoder excitation with dynamically controlled voicedness. AFCRL and Institute of Electrical and Electronic Engineers Audio and Electro-acoustics 1967 Conference on Speech Communication and Processing. Bedford: OAR, United States Air Force, 1967.

Kelly, J., & Miller, R. Recent improvements in 4800 BPS voice-excited vocoders. AFCRL and Institute of Electrical and Electronic Engineers Audio and Electro-acoustics 1967 Conference on Speech Communication and Processing. Bedford: OAR, United States Air Force, 1967.

Kimura, D. Cerebral dominance and the perception of verbal stimuli. Canadian Journal of Psychology, 1961, 15(3), 166-171. (a)

Kimura, D. Some effects of temporal-lobe damage on auditory perception. Canadian Journal of Psychology, 1961, 15(3), 156-165. (b)

Kimura, D. Speech lateralization in young children as determined by an auditory test. Journal of Comparative and Physiological Psychology, 1963, 56(5), 899-902.

Kimura, D. Left-right differences in the perception of melodies. Journal of Experimental Psychology, 1964, 16(4), 355-358.

Klare, G. R. The measurement of readability. Ames, Iowa: Iowa State University Press, 1963.

Klineman, J. The effects of training session on the ability to comprehend compressed speech. Unpublished manuscript, University of Pittsburgh, 1963.

Klumpp, R. G., & Webster, J. C. Intelligibility of time-compressed speech. Journal of the Acoustical Society of America, 1961, 31, 265-267.

Kodman, F., Jr. An investigation of word intelligibility as a function of time compression. Unpublished doctoral dissertation, University of Illinois, 1954.

Koen, F., Becker, A., & Young, R. The psychological reality of the paragraph. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 49-53.

Kozhevnikov, V. A., & Chistovich, L. A. Speech: Articulation and perception (Moscow-Leningrad). Translated by J. P. R. S., Washington, D. C., No. JPRS 30, 543, 1965.

Kryter, K. On predicting the intelligibility of speech from acoustic measures. Journal of the Acoustical Society of America, 1956, 28.

Kryter, K. Speech bandwidth compression through spectrum selection. Journal of the Acoustical Society of America, 1960, 32, 547-556.

Kuchenmeister, N. V. Internationale Maatschappij voor Sprekende Films. Improvements relating to the reproduction of sounds from records. London: The Patent Office, January 26, 1933. British Patent No. 386,983.

Kurtzrock, G. H. The effects of time and frequency distortion upon word intelligibility. Speech Monographs, 1957, 24, 94.

LaBenz, P. Potentialities of auditory perception for various levels of hearing loss. Volta Review, 1956, 58, 397-402.

Lacy, O. L., & Siegel, P. S. Analysis of the unit of measurement of the galvanic skin response. Journal of Experimental Psychology, 1949, 39, 122-127.

Lafon, J. C. The phonetic test and the measurement of hearing. Springfield: C. C. Thomas, 1966.

Lafon, J. C. Appareil amplificateur a compensation pur la sourdite profonde. Journal of International Audiology, 1967, 6, 343-345.

Lass, N. J. A comparative study of rate characteristics in cleft palate and non-cleft palate speakers. Unpublished doctoral dissertation, Purdue University, 1968.

Laughery, K., Lachman, R., & Dansereau, D. Short-term memory: Effects of item-pronunciation time. In, Proceedings of the 73rd annual convention of the American Psychological Association. Washington, D. C.: American Psychological Association, 1965.

Laughery, K., & Pinkus, A. Recoding and presentation rate in short-term memory. Journal of Experimental Psychology, 1968, 76, 636-641.

Lawton, R. Listening to rate-controlled speech: A technique for improving education and communication in the Air Force and industry. Unpublished master's thesis, George Washington College, 1967. Also in: Unpublished thesis, Air Command and Staff College, Air University, Maxwell Air Force Base, 1967.

Leitner, M. A. A study of the effects of intraphase rate and pause time on information gain and speaker image. (Unpublished doctoral dissertation, University of Wisconsin, 1962) Dissertation Abstracts, 1962, 23, 2251.

Lenneberg, E. H. On explaining language. Science, 1969, 164, 635-643.

Liberman, A. M. Some results of research on speech perception. Journal of the Acoustical Society of America, 1957, 29, 117-123.

Liberman, A. M. Some results on speech perception. In S. Sapolta (Ed.), Psycholinguistics. New York: Holt, Rinehart & Winston, 1961. (Pp. 142-153)

Library of Congress. Report on compressed speech activities at the Division for the Blind. Washington, D. C.: Library of Congress, 1966.

Licklider, J. The intelligibility of amplitude-dichotimized, time-quantized speech waves. Journal of the Acoustical Society of America, 1950, 22, 677.

Lieberman, P. Intonation, perception and language. Cambridge: The MIT Press, 1967.

Lindquist, E. P. Design and analysis of experiments in education and psychology. Boston: Houghton Mifflin, 1953.

Ling, D. The use of low frequency residual hearing in profoundly deaf children. Unpublished master's thesis, McGill University, 1966.

Ling, D. Auditory discrimination of coded speech by deaf children. Unpublished doctoral dissertation, McGill University, 1968. (a)

Ling, D. Transposition defrequences et production des mots. (Frequency transposition and speech production.) Ami Sourds, 1968, 9-11. (b)

Ling, D. Three experiments on frequency transposition. American Annals of the Deaf, 1968, 113, 283-294. (c)

Ling, D., & Doehring, D. G. Learning limits of deaf children for coded speech. Journal of Speech and Hearing Research, 1969, 12, 83-94.

Ling, D., & Druz, W. Transposition of high frequency sounds by partial vocoding of the speech spectrum: Its use by deaf children. Journal of Auditory Research, 1967, 7, 133-144.

Loper, J. L. An experimental study of some effects of time compression upon the comprehension and retention of a visually augmented televised speech. Unpublished doctoral dissertation, University of Southern California, 1966.

Lovitt, T. Operant preference of retarded and normal males for rate of narration. Psychological Review, 1968, 18, 205-214.

Lown, A. Some observations on speeded speech by visually-impaired students in the Atlantic public schools. In. Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967.

Lutterman, D. M., Welsh, O. L., & Melrose, J. Responses of aged males to time-altered speech stimuli. Journal of Speech and Hearing Research, 1966, 9, 226-230.

MacKay, D. Metamorphosis of a critical interval: Age-linked changes in the delay in auditory feedback that produces maximal disruption of speech. Journal of the Acoustical Society of America, 1968, 43, 625-627.

Maclay, H., & Osgood, C. E. Hesitation phenomena in a spontaneous English speech. Word, 1959, 15, 19-44.

MacLean, D. Analysis of speech in a helium-oxygen mixture under pressure. Journal of the Acoustical Society of America, 1966, 40, 625-627.

Mahaffey, R., & McDill, J. Some relations between perception of short silent intervals and the production of speech. Journal of the Acoustical Society of America, 1968, 44, 396.

Malecot, A. Un segmenteur-explorateur de bandes magnetiques. Rev. Gen. L'enseign. Deficient Auditifs, 1968, 60, 81-85.

Markgraf, B. R. Demands on the listening skills of secondary school pupils. In S. Duker (Ed.), *Listening: Readings*. Metuchen, N. J.: Scarecrow Press, 1966. (Pp. 90-94)

Martin, D., Murphy, R., & Meyer, A. Articulation reduction by combined distortions of speech waves. *Journal of the Acoustical Society of America*, 1956, *28*, 597-601.

Martin, J. Temporal word spacing and the perception of ordinary, anomalous, and scrambled strings. *Journal of Verbal Learning and Verbal Behavior*, 1968, *7*.

Martony, J., & Pickett, J. Low frequency vowel formant discrimination in deaf listeners. Unpublished manuscript, Gallaudet College, 1969.

McBroom, P. Compressed speech helps. *Science Newsletter*, 1965, *88*, 214.

McCarthy, J. J., & Kirk, S. A. *The Illinois test of psycholinguistic abilities, experimental edition*. Urbana, Ill.: University of Illinois Institute for Research on Exceptional Children, 1961.

McLain, J. A comparison of two methods of producing rapid speech. *International Journal for the Education of the Blind*, 1962, *12*(2), 40-43.

McNemar, Q. *Psychological statistics*. New York: Wiley, 1962.

Miller, E. C. Effects on learning of variations in oral presentation. Unpublished doctoral dissertation, University of Denver, 1954.

Miller, G. A. Intelligibility of speech: Effects of distortion. In C. E. Waring (Ed.), *Combat Instrumentation II*. Washington, D. C.: Government Printing Office, 1946. (Pp. 88-109)

Miller, G. A. *Language and communication*. New York: McGraw-Hill, 1951. (a)

Miller, G. A. Speech and language. In S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. (b)

Miller, G. A. What is information measurement? *American Psychologist*, 1953, *8*, 3-11.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, *63*, 81-97.

Miller, G. A. Decision units in the perception of speech. *Transactions of the Institute of Radio Engineers Professional Group on Information Theory*, 1962, *IT-8*, 81-83.

Miller, G. A., & Chomsky, N. Finitary models of language users. In (R. Luce, R. Bush, and E. Galanter, Eds.), *Handbook of mathematical psychology*, 1963, *2*, 419-492.

Miller, G. A., Heise, G. A., & Lichten, W. The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 1951, *41*, 329-335.

Miller, G. A., & Isard, S. Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, 1963, *2*, 217-228.

Miller, G. A., & Licklider, J. C. R. The intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, 1950, *22*, 167-173.

Miller, G. A., & Selfridge, J. A. Verbal context and the recall of meaningful material. *American Journal of Psychology*, 1950, *63*, 176-185.

Miller, R. L. *Signaling system*. Washington, D. C.: United States Patent Office, May 17, 1938. U. S. Patent No. 2,117,739.

Minifie, F. D. An analysis of the durational aspects of connected speech by means of an electronic speech duration analyzer. Unpublished doctoral dissertation, University of Iowa, 1963.

Miron, M. S. (Ed.) *Experimental phonetics: Selected articles by Grant Fairbanks*. Urbana, Ill.: University of Illinois Press, 1966.

Miron, M. S., & Brown, E. R. Stimulus parameters in speech compression. *Journal of Communication*, 1968, *18*, 219-235.

Moray, N. *Attention: Selective processes in hearing and vision*. London: Hutchinson, 1969.

Murdock, B. Recent developments in short-term memory. *British Journal of Psychology*, 1967, *3/4*, 421-433.

Murray, D. Visual and auditory presentation, presentation rate, and short-term memory in children. British Journal of Psychology, 1968, 59, 119-125.

Neisser, U. Cognitive psychology. New York: Appleton, 1967.

Nelson, H. E. The effect of variations of rates on the recall by radio listeners of straight newscasts. Speech Monographs, 1948, 15, 173-180.

Nelson, M. J., & Denny, E. C. The Nelson-Denny reading test, revised by J. I. Brown. Boston: Houghton Mifflin, 1960.

Nichols, R. G., & Stevens, L. A. Are you listening? New York: McGraw-Hill, 1957.

Nixon, C. W., Mabson, W. E., Trimboll, F., Endicott, J. E., & Welch, B. E. Observations on man in an oxygen-helium environment at 380 mm. Hg total pressure: IV. Communications. Aerospace Medicine, 1968, 39(1), 1-9.

Nixon, C. W., & Sommer, H. C. Subjective analysis of speech in helium environments. Aerospace Medicine, 1968, 39(2), 139-144.

Nober, E., & Nober, L. Speech reception thresholds and discrimination scores as a function of method of presentation and frequency response. Journal of Auditory Research, 1962, 2, 1-5.

Nolan, C. Y. Reading and listening in learning by the blind. (Progress Report, PHS Grant No. NB-04870-04) Louisville: American Printing House for the Blind, 1966.

Nolan, C. Y. Listening and reading in learning. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967.

Nolan, C. Y. Reading and listening in learning by the blind. (Terminal Progress Report, PHS Grant No. NB-048-70) Louisville: American Printing House for the Blind, 1968.

Norman, D. Memory and attention: An introduction into human information processing. New York: Wiley & Sons, 1969.

Ochai, Y. General considerations on studies of speech qualities in rotational synchronous distortion. Mem. Fac. Eng. Nagoya, 1955, 7, 36-39.

Ochai, Y., & Fukumura, T. Study on fundamental qualities of vocalic timbre by rotational synchronous distortion. Mem. Fac. Eng. Nagoya, 1956, 8, 1-10.

Ochai, Y., & Imumitachi, H. Timbre study of mishearing phonemes of speech phones in rotational synchronous distortion, Report I: Mishearing of vowel timbre. Mem. Fac. Eng. Nagoya, 1955, 7, 49-60.

Ochai, Y., Saito, S., & Watambe, Y. Allowance problem in rotational synchronous distortion as a study on timbre discrimination by infinitesimal position-shift in the so-called timbre-space. Mem. Fac. Eng. Nagoya, 1955, 7, 131-144.

Oeken, P. Can the hearing of patients suffering from high tone perceptive deafness be improved by frequency transposition? Journal of International Audiology, 1963, 2, 263-266.

Orr, D. B. Note on thought rate as a function of reading and listening rates. Perceptual and Motor Skills, 1964, 19, 874.

Orr, D. B. Recent research on reading and the comprehension of time-compressed speech. Proceedings of the College Reading Association, Easton, Pennsylvania, 1966, 7, 79-84.

Orr, D. B. Preface to the special edition of the Journal of Communication on speech compression. Journal of Communication, 1968, 18, 195-197. (a)

Orr, D. B. Time-compressed speech--a perspective. Journal of Communication, 1968, 18, 288-292. (b)

Orr, D. B., & Friedman, H. L. Research on speeded speech as an educational medium. (Progress Report Grant #7-48-7670-203) Washington, D. C.: United States Department of Health, Education, and Welfare, Office of Education, 1964.

Orr, D. B., & Friedman, H. L. The effect of listening aids on the comprehension of time-compressed speech. Journal of Communication, 1967, 17, 223-227.

Orr, D. B., & Friedman, H. L. Effects of massed practice on the comprehension of time-compressed speech. Journal of Educational Psychology, 1968, 59, 6-11.

Orr, D. B., Friedman, H. L., & Williams, J. C. C. Trainability of listening comprehension of speeded discourse. Journal of Educational Psychology, 1965, 56(3), 148-156.

Osgood, C. E. Method and theory in experimental psychology. New York: Oxford University Press, 1953.

Osgood, S., Suci, G., & Tannenbaum, P. The measurement of meaning. Urbana, Ill.: University of Illinois Press, 1957.

Otis, A. S. Otis group intelligence test. New York: Harcourt, Brace & World, 1918/1940.

Otis, A. S. Otis quick-scoring mental ability tests. New York: Harcourt, Brace & World, 1954.

Overmann, R. A. Processing time as a variable in the comprehension of time-compressed speech. Unpublished master's thesis, University of Louisville, 1969. Also in: Proceedings of the Second Louisville Conference on Rate and/or Frequency-Controlled Speech. Louisville: University of Louisville, Perceptual Alternatives Laboratory, 1970.

Peters, R. W. The effect of changes in side-tone delay and level upon rate of oral reading of normal speakers. Journal of Speech and Hearing Disorders, 1954, 19, 483-490.

Peterson, G. The significance of various portions of the wavelength in the minimum duration necessary for the recognition of vowel sounds. Unpublished doctoral dissertation, Louisiana State University, 1939.

Peterson, G., Sivertsen, E., & Subramanyam, D. Intelligibility of diphasic speech. Journal of the Acoustical Society of America, 1956, 28, 405-411.

Peterson, H. The study of normal hearing childrens' discrimination for two types of verbal materials presented with varying degrees of distortion. Unpublished doctoral dissertation, University of Illinois, 1967.

Pickett, J. Conference on speech analyzing aids for the deaf: Historical notes and preface. American Annals of the Deaf, 1967, 113, 117-119.

Pickett, J. Recent research on speech analyzing aids for the deaf. Institute of Electrical and Electronic Engineers Transactions on Audio and Electro-acoustics, 1968, AU-15, 227-234. (a)

Pickett, J. Sound patterns of speech. American Annals of the Deaf, 1968, 113, 120-126. (b)

Some comparative measurements of
ı spectral differences. American
9-267.

Intelligibility of excerpts from fluent
ce and duration of excerpt. Language

ı de la parole synthetique dans la cor-
12, 285-290. (English translation in:
53, 3, 73-82.)

ıthetique et son application dans la cor-
stique, 1965, 133/134, 151-187.

bility and message reception. Journal
ica, 1964, 36, 937-945.

Intelligibility of excerpts from fluent
al context. Journal of Verbal Learning
-84.

, & Decker, L. Analysis of incorrect
set. Journal of the Acoustical Society

A. H. (Eds.) Perspectives.
1963.

R. F. Chronometric analysis of
iew, 1967, 74, 392-409.

E., & Silverstein, A. The running
y Journal of Experimental Psychology,

Toward the specification of speech.
of America, 1950, 22, 807-820.

. In A. Sovijarvi & P. Aaito (Eds.),
ional congress of phonetical sciences.

ıemory. Air Force Cambridge Research
No. 8668. Cambridge, Mass.: Bolt,

epts: A theory and simulation of some
avioral Science, 1967, 12, 412-430.

Quillian, M. R. The teac
lation program and theory of lar
tion for Computing Machinery,

Qureshi, S. U. H., & Kin
speech on a small computer. In
conference on rate and/or frequ
University of Louisville, Percep

Ramsay, R. W., & Law,
of speech. *Language and Speec*

Raymond, T., & Proud, C
of Otolaryngology, 1962, *76*, 43

Reddy, D. R. Segmentati
Acoustical Society of America,

Reid, R. H. Grammatica
pressed speech. *Journal of Con*

Reidford, P. Reading con
of speech compression training.
sity of Illinois, 1965.

Resta, P. E. The effects
comprehension of frequency shi
blind. AEEM-298, Goodyear Ae
Arizona, May 15, 1968.

Reynolds, L. Auditory co
Unpublished master's thesis, Ur

Risberg, A. A bibliograp
and hard of hearing. *Speech Tra*
Status Report, 1964, 21-30.

Risberg, A. The 'transpo
Speech Trans. Laboratory Quart
26-30.

Robins, J. E. A study of
immediate recall of speeded spe
dissertation, George Washington

Robinson, J. A. Category clustering: A "plan" for category retrieval in free recall. Technical Report No. 1 for AF-AFOSR-1008-66. Louisville: University of Louisville, 1966. (a)

Robinson, J. A. Category clustering in free recall. *Journal of Psychology*, 1966, *62*, 279-285. (b)

Robinson, J. A. Scanning and clustering: An exploratory investigation of retrieval strategies in free recall. Technical Report No. 2 for AF-AFOSR-1008-66. Louisville: University of Louisville, 1966. (c)

Rodgers, J. R. A formula for predicting the comprehension level of material to be presented orally. *Journal of Educational Research*, 1962, *56*, 218-220.

Rose, H. Performance evaluation of clipped speech. *AFCRL and Institute of Electrical and Electronic Engineers Audio and Electro-acoustics 1967 Conference on Speech Communication and Processing*. Bedford: OAR, United States Air Force, 1967.

Rothauser, E., & Maiwald, D. Simulated vocoder analysis. *Journal of the Acoustical Society of America*, 1968, *44*, 390.

Rothauser, E., Urbansk, G., & Pachl, A. A comparison of preference measurement methods. *AFCRL and Institute of Electrical and Electronic Engineers Audio and Electro-acoustics 1967 Conference on Speech Communication and Processing*. Bedford: OAR, United States Air Force, 1967.

Rupf, J., Hughes, G., & House, A. Interaural switching effects on the recognition of consonants. Paper presented at the 77th Meeting of the Acoustical Society of America, Philadelphia, 1969.

Savin, H. B. Word-frequency effect and errors in the perception of speech. *Journal of the Acoustical Society of America*, 1963, *35*, 200-206.

Scharf, D. Intelligibility of reiterated speech. *Journal of the Acoustical Society of America*, 1959, *31*, 423-427.

Schwartz, E., & Kleiboener, A. A language element for compression coding. *Information and Control*, 1967, *10*, 315-333.

Scheisser, H. A device for time expansion used in sound recordings. *Transactions of the Institute for Radio Engineers Professional Group on Audio*, 1954, *AU-2*, 12-15.

Schimmel, H., & Clay, S. C., Jr. Variable-delay, rotating, signal-delay, and time compression device. Washington, D. C.: United States Patent Office, October 8, 1963. U. S. Patent No. 3,106,705.

Scholes, R. J. On the spoken disambiguation of superficially ambiguous sentences. Bell Telephone Laboratories Technical Memorandum, July, 1968.

Scholtz, P. N., & Bakis, P. Spoken digit recognition using vowel-consonant segmentation. Journal of the Acoustical Society of America, 1962, 34, 1-5.

Schroeder, M. Vocoders: Analysis and synthesis of speech. Proceedings of the Institute of Electrical and Electronic Engineers, 1966, AU-54, 720-734.

Schuller, E. Horkipf zum abtasten von magnetogrammen mit gegenuber der aufnahmegeschwindigkeit veranderter wiedergabegeschindigkeit. German Patent Office, April 23, 1942. Patent No. 721,198.

Schuller, E. Sound reproducer. Washington, D. C.: United States Patent Office, June 20, 1944. U. S. Patent No. 2,352,023.

Schultz, T. The economic value of education. New York: Columbia University Press, 1963.

Schwartz, M. F. Differential effect of instructions upon the rate of oral reading. Journal of the Acoustical Society of America, 1961, 33, 1801-1802.

Scott, R. J. Temporal effects in speech analysis and synthesis. Unpublished doctoral dissertation, University of Michigan, 1965. Also in: Communication Scientist Laboratory, Contract No. NONR 1224 (22), NR 049-122, 1965.

Scott, R. J. Computer for speech time compression. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967. (a)

Scott, R. J. Time adjustment in speech synthesis. Journal of the Acoustical Society of America, 1967, 41, 60-65. (b)

Seo, H. New approach to speech compression. Journal of the Acoustical Society of America, 1967, 41, 1595. (a)

Seo, H. Speech compression. Unpublished doctoral dissertation, University of South Carolina, 1967. (b)

Seo, H. Electronic speech compression. Champaign, Ill.: Bioacoustics Laboratory, University of Illinois, 1969.

Sequential tests of educational progress: Listening. Princeton, N. J.: Educational Testing Service, Cooperative Test Division, 1963.

Sergeant, R. Phonemic analysis of consonants in helium speech. Journal of the Acoustical Society of America, 1967, 41, 66-69.

Shearn, D., Sprague, R., & Rosenzweig, S. A method for the analysis and control of speech rate. Journal of the Experimental Analysis of Behavior, 1961, 4, 197-201.

Shriner, T., Beasley, D., & Zemlin, W. Effects of frequency-divided speech signals on identification accuracy and reaction time measures. Journal of Speech and Hearing Research, 1970, in press.

Shriner, T., & Daniloff, R. Reassembly of segmented CVC syllables. Journal of Speech and Hearing Research, 1970, in press.

Shriner, T., Daniloff, R., & Nemac, J. Resynthesization of meaningful and non-meaningful CVC syllables. Champaign, Ill.: University of Illinois, 1969.

Shriner, T., & Sprague, R. The effects of time compression on speech identification accuracy and reaction time measures in normals versus retardates. Champaign, Ill.: University of Illinois, 1969.

Shriner, T., & Sprague, R. The effects of time-compressed speech signals on identification accuracy and latency measures. Journal of Experimental Child Psychology, 1970, in press.

Siegel, C. M. Interexaminer reliability for mean length of response. Journal of Speech and Hearing Research, 1962, 5, 91-95.

Siegel, S. Nonparametric statistics. New York: McGraw-Hill, 1956.

Silverman, F. An approach to determining the number of judges needed for scaling experiments. Perceptual Motor Skills, 1968, 27, 1333-1334.

Simpson, E. A. Better reading book 2. Chicago: Science Research Associates, 1963.

Sitter, T. Short term retention of sequentially presented digits as a function of interdigit interval, digit duration, and series length. Journal of Experimental Psychology, 1968, 78, 174-178.

Slawson, A. Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency. Journal of the Acoustical Society of America, 1968, 43, 87-101.

Slosson, R. Slosson oral reading tests. East Aurora, N. Y.: Educational Publications, 1963.

Snedecor, G. W., & Cochran, W. G. Statistical methods (6th Ed.). Ames, Iowa: The Iowa State University Press, 1967.

Snidecor, J. C. A comparative study of the pitch and duration characteristics of impromptu speaking and oral reading. Speech Monographs, 1943, 10, 50-56.

Speaks, C. Intelligibility of filtered synthetic sentences. Journal of Speech and Hearing Research, 1967, 10, 289-298. (a)

Speaks, C. Performance-intensity characteristics of selected verbal materials. Journal of Speech and Hearing Research, 1967, 10, 344-347. (b)

Speaks, C. Estimating the intelligibility of continuous discourse. Paper presented at the 77th Meeting of the Acoustical Society of America, Philadelphia, 1969.

Speaks, C., & Jerger, J. Method for measurement for speech identification. Journal of Speech and Hearing Research, 1965, 8, 185-194.

Speaks, C., & Jerger, J. Synthetic sentence identification and the receiver operating characteristics. Journal of Speech and Hearing Research, 1967, 10, 110-119.

Speaks, C., & Karmen, J. The effect of noise on synthetic sentence identification. Journal of Speech and Hearing Research, 1967, 10, 859-864.

Speaks, C., Karmen, J., & Benitz, L. Effect of a competing message on synthetic speech identification. Journal of Speech and Hearing Research, 1967, 10, 390-395.

Speech compression 1953 to the present. Arlington Armed Services Technical Information Agency (AD 2888 50), 1961.

Spicker, H. Listening comprehension and retention of normal and retarded children as functions of speaking rate and passage difficulty. (Unpublished doctoral dissertation, George Peabody College, 1963) Dissertation Abstracts, 1964, 24, 1925.

Springer, A. M. Rotatable electromagnetic transducers. Washington, D. C.: United States Patent Office, March 28, 1961. U. S. Patent No. 2,977,423. (a)

Springer, A. M. Variable drive friction controlled recording system. Washington, D. C.: United States Patent Office, August 15, 1961. U. S. Patent No. 2,996,583. (b)

Springer, A. M. Electromagnetic transducer. Washington, D. C.: United States Patent Office, November 13, 1962. U. S. Patent No. 3,064,088. (a)

Springer, A. M. Magnetic recording and playback system for magnetic tape and magnetic wire. Washington, D. C.: United States Patent Office, June 26, 1962. U. S. Patent No. 3,041,000. (b)

Springer, A. M. Rotatable magnetic systems for transforming audio currents. Washington, D. C.: United States Patent Office, July 31, 1962. U. S. Patent No. 3,047,673. (c)

Springer, A. M. Magnetic recording and play-back systems. Washington, D. C.: United States Patent Office, February 12, 1963. U. S. Patent No. 3,077,587.

Staats, C. K., Staats, A. W., & Schutz, R. E. The effects of discrimination pretraining on textural behavior. Journal of Educational Psychology, 1962, 33, 32-37.

Starkweather, J. A. A speech rate meter for vocal behavior analysis. Journal of the Experimental Analysis of Behavior, 1960, 3, 111-114.

Starkweather, J. A., & Hargreaves, W. A. The influence of sodium pentobarbital on vocal behavior. Journal of Abnormal Social Psychology, 1964, 69, 123-126.

Stauffer, R. G., Burrows, A. T., & Jones, D. M. Skyways to tomorrow. New York: Holt, Rinehart & Winston, 1962.

Steinberg, J. D. Acoustics, the sense of hearing. In H. Pender & K. McIlwain (Eds.), Electrical engineer's handbook. New York: John Wiley & Sons, 1936. (Pp. 9-32)

Stevens, H. A., & Heber, R. (Eds.) Mental retardation: A review of research. Chicago: University of Chicago Press, 1964.

Stevens, K. Review of existing speech compression systems. Griffith Air Force Base, New York: Rome Air Development Center, 1961.

Sticht, T. Some relationships of mental aptitude, reading ability, and listening ability using normal and time-compressed speech. Journal of Communication, 1968, 18, 243-258.

Sticht, T. Some interactions of speech rate, signal distortion, and certain linguistic factors in listening comprehension. Audio-Visual Communication Review, 1969, 17(2), 159-171.

Sticht, T., & Gray, B. B. Intelligibility of time-compressed words as a function of age and hearing loss. Journal of Speech and Hearing Research, 1969, 12, 443-448.

Stone, S. Prior entry in the auditory-tactual complication. American Journal of Psychology, 1926, 37, 284-287.

Stover, W. Electronic speech processing methods for the hearing impaired. Journal of Auditory Research, 1967, 7, 313-325. (a)

Stover, W. Technique for correcting helium speech distortion. Journal of Acoustical Society of America, 1967, 41, 70-74. (b)

Stover, W. Time-domain bandwidth-compression system. ournal of the Acoustical Society of America, 1967, 42, 348-359. (c)

Takefuta, Y., & Swigart, E. Intelligibility of speech signals spectrally compressed by a sampling-synthesizing technique. Institute of Electrical and Electronic Engineers Transactions on Audio and Electroacoustics, 1968, AU-15, 271-274.

Taub, H., Monty, R., & Laughery, K. Keeping track of sequential events: Effects of stimulus on-time and interstimulus off-time. United States Army Technical Memorandum (Human Engineering Laboratory), 1967, 6-67.

Taylor, E. A. *The fundamental reading skill, as related to eye movement photography and visual anomalies*. (2nd ed.) Springfield, Ill.: Thomas, 1966.

Taylor, J. Use of compressed speech, tapes, and discs of variable frequency power supply with selected children and adults. In, *Proceedings of the Louisville conference on time-compressed speech*. Louisville: University of Louisville, 1967.

Taylor, S. E., Frankenpohl, H., et al. (Eds.) *Listen and read*. Huntington, N. Y.: Educational Development Laboratory, 1962.

Templin, M. C. *Certain language skills in children, their development and interrelationship*. Minneapolis: University of Minnesota Press, 1957.

Terengo, L. Pitch and duration characteristics of the oral reading of males on a masculinity-feminity dimension. *Journal of Speech and Hearing Research*, 1966, *9*, 590-595.

Thomas, I. A second formant vocoder. *AFCRL and Institute of Electrical and Electronic Engineers Audio and Electro-acoustics 1967 Conference on Speech Communication and Processing*. Bedford: OAR, United States Air Force, 1967. (a)

Thomas, I. Speech synthesis and recognition. *Biological Computer Laboratory Report No. 67.2*. Champaign: University of Illinois, Biological Computer Laboratory, 1967. (b)

Thomas, I. The significance of the second formant in speech intelligibility. Biological Computer Laboratory Technical Report No. 10, Eng. Res. Lab., Eng. Exp. Station. Champaign: University of Illinois, 1968. (a)

Thomas, I. Structural and dynamic aspects of speech signals. *Biological Computer Laboratory Report No. 682*. Champaign: University of Illinois, 1968. (b)

Thomas, I., & Flavin, R. The intelligibility of speech transposed downward in frequency by one octave. *Journal of AFS*, February 1969.

Thurstone, L. L., & Thurstone, T. G. *Examiners' manual, PMA, primary mental abilities for grades 4-6*. Chicago: Science Research Associates, 1963.

Tiffany, W., & Bennett, D. The intelligibility of slow-played speech. *Journal of Speech and Hearing Research*, 1961, *4*, 248-258.

Travers, R. M. W. The transmission of information to human receivers. Audio-Visual Communication Review, 1964, 12, 373-385.

Travers, R. M. W. Studies related to the design of audiovisual teaching materials. (Final Report, Contract No. OEC 3-20-003) Washington, D. C.: United States Department of Health, Education, and Welfare, Office of Education, 1966.

Travis, L. E., & Rasmus, B. The speech sound discrimination ability of cases with functional disorders of communication. Quarterly Journal of Speech, 1931, 17, 217-226.

Treisman, A. M. The effects of redundancy and familiarity on translating and repeating back a foreign and a native language. British Journal of Psychology, 1965, 56, 369-379.

Vilbig, F. An apparatus for speech compression and expansion and for replaying visible speech records. Journal of the Acoustical Society of America, 1950, 22, 754-761.

Vilbig, F. Frequency-bank multiplication or division and time expansion or compression by means of a string filter. Journal of the Acoustical Society of America, 1954, 24, 33-39.

Vilbig, F. Frequency band compression and expansion by respective exclusion and repetition of time segments. (Report No. TR 57-104) Bedford: AFCRL, 1957.

Vilbig, F. Speech frequency band compression and speech redundancy-reduction by means of vocoders. Ottobrunn bei Munchen. Germany: Bolkow GMBH, 1967.

Vilbig, F., & Haase, K. Some systems of speech band compression. Journal of the Acoustical Society of America, 1956, 28, 573-577.

Voiers, W. On the problem of reliability, sensitivity, and diagnostic value in tests of speech intelligibility. AFCRL and Institute of Electrical and Electronic Engineers Audio and Electro-acoustics 1967 Conference on Speech Communication and Processing. Bedford: OAR, United States Air Force, 1967.

Voiers, W. The present state of digital vocoding technique: A diagnostic evaluation. Institute of Electrical and Electronic Engineers Transactions on Audio and Electro-acoustics, 1968, AU-15, 275-279.

Volz, H. Frequenztransponierung durch rotierende magne-Honkopfe. Bild und Ton., 1961, 14, 384-390.

Von Oeken, F. Investigations into the intelligibility of speech and recognizability of phonemes after frequency transposition. Phonetica, 1963, 110, 92-107.

Voor, J. B. The effect of practice upon the comprehension of time-compressed speech. Unpublished master's thesis, University of Louisville, 1962.

Voor J. B., & Miller, J. The effect of practice upon the comprehension of time-compressed speech. Speech Monographs, 1965, 32, 452-455.

Walker, C. Functional units in speech perception, as disclosed by time-compressed speech. Unpublished doctoral dissertation (in progress), University of Illinois, 1969.

Walker, H. W., & Lev, J. Statistical inferences. New York: Holt, Rinehart & Winston, 1953.

Wang, W. S. Y., & Fillmore, C. Intrinsic cues and consonant perception. Journal of Speech and Hearing Research, 1961, 4, 130-136.

Wason, P. C. The retention of material presented through precis. Journal of Communication, 1962, 12, 36-43.

Webster, J., Davis, H., & Ward, W. Everyday speech intelligibility. Journal of the Acoustical Society of America, 1965, 38, 668.

Wechsler, D. Wechsler adult intelligence scale. New York: Psychological Corporation, 1949.

Wedenburg, E. Auditory training of the severely hard of hearing using a coding amplifier. Proceedings of the 3rd International Congress of Acoustics. Amsterdam: Elsevier, 1961. (Pp. 657-660)

Weiss, D. A. Cluttering. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Welford, A. T. Fundamentals of skill. London: Methuen, 1968.

Wenzel, R. J. Speech compression studies. Unpublished master's thesis, Massachusetts Institute of Technology, 1962.

Western Electric Co., Ltd. Improvements in wave signalling systems. London: British Patent Office, October 14, 1926. Patent No. 259,328.

Wilkes, A., & Kennedy, R. Relationship between pausing and retrieval latency in sentences of varying grammatical form. Journal of Experimental Psychology, 1969, 79, 241-245.

Williams, C., & Hecker, M. Relation between intelligibility scores for test methods and three types of speech distortion. Journal of the Acoustical Society of America, 1968, 44, 1002-1006.

Williams, E. J. Regression analysis. New York: Wiley & Sons, 1959.

Wilson, J. Effects of stimulus and interstimulus duration on the immediate recall of time-compressed sequences of different orders of approximation to English. Unpublished master's thesis, University of Louisville, 1969.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Winges, S. A. The intelligibility of interrupted speech. Unpublished master's thesis, University of Louisville, 1963.

Wood, C. D. Comprehension of compressed speech by elementary school children. (Final Progress Report Number 1300, Grant No. 7-24-0210-263) Washington, D. C.: United States Department of Health, Education, and Welfare, Office of Education, 1965.

Woodcock, R., & Clark, C. Comprehension of a narrative passage by elementary school children as a function of listening rate, retention period, and IQ. Journal of Communication, 1968, 18(3), 259-271. (a)

Woodcock, R., & Clark, C. R. Comprehension of a narrative passage by fifth-grade children as a function of listening rate and presentation strategy. Institute on Mental Retardation and Intellectual Development Papers and Reports, 5(3), 1968. (b)

Woodcock, R., & Clark, C. R. Differences in learning through compressed speech as a function of presentation strategy and rate among culturally disadvantaged fourth-grade children. Institute on Mental Retardation and Intellectual Development Papers and Reports, 5(9), 1968. (c)

Woodcock, R., & Clark, C. R. Influence of presentation rate and media on the comprehension of narrative material by adolescent educable retardates. Institute on Mental Retardation and Intellectual Development Papers and Reports, 5(7), 1968. (d)

Woodworth, R. S., & Schlosberg, H. Experimental psychology. New York: Henry Holt & Co., 1954. (P. 33)

Wurm, S. Pitch and intensity recording devices for the study of Australasian languages. Zeitschrift fur Phonetick, 1967, 20, 250-257.

Zemlin, W. The use of bandwidth and time compression for the hard of hearing handicapped. In, Proceedings of the Louisville conference on time-compressed speech. Louisville: University of Louisville, 1967.

Zemlin, W., Daniloff, R., & Shriner, T. The difficulty of listening to time-compressed speech. Journal of Speech and Hearing Research, 1968, 11, 673-896.

Zinnenberg, M. On the study and retraining of hearing in children possessing a severe hearing loss. In R. W. West (Ed.), American Speech and Hearing Reports, 1968, 3, 140-146.